BANGALORE (India)

December 8-10, 1993

INTERNATIONAL 93 BANGALORE CONFERENCE

SIGNALS, DATA, SYSTEMS

SDS'93

◆

PROCEEDINGS

Vol. 1

◆

Organized by

AMSE, Association for the Advancement of
Modelling and Simulation techniques in Enterprises

ISBN: 2-909214-41-9

PROCEEDINGS

INTERNATIONAL AMSE CONFERENCE
SIGNALS, DATA, SYSTEMS
SDS'93

Bangalore (India), December 8-10, 1993

Vol. 1: Methodologies
Socio-economic problems

TABLE

p. 3 S. GHOSH- *On characterisation of a distribution function in terms of the value of its characteristic function at some points*

15 S. GHOSH- *Sharpening of a theorem of Gwo Dong Lin and an application*

25 S. GHOSH- *A set of necessary and sufficient conditions for perfect mixture of perfect measures being perfect*

39 S. GHOSH- *On gamma-functions and sum of independent random variables*

53 S. GHOSH- *A probabilistic approach to Stone-Weierstrass approximation theorem*

67 K.J. SATAO- *Development of 'Rename Directory' command in MS-DOS*

79 K.J. SATAO- *Development of 'Search' and 'Replace' commands in MS-DOS using a linked list method*

91 TESSAMMA THOMAS, C.S. SRIDHAR- *A phoneme identification system*

101 K.S. JOG, S.S. BHAT, M.Y. NAGARWALA- *Adaptive autoregressive modelling for restoration of music*

111 M. MEENAKSHI, V.C. RAVICHANDRAN- *On the Gaussian approximation analysis of m-ary signalling in direct detection optical CDMA communication systems*

135 R. PANDA, S. PATNAIK, G.S. RATH- *Fuzzy entropy net: Application to recognition of handprinted Oriya characters*

147 M.A. AL RASHEEDI, A. RAOUF MOHAMED, A. MANSOUR, M.G. ALI- *Modified window function for finite-impulse-response digital filters and data detection*

155 K.A. GOPALA RAO, N. VISWANADHAM- *A new approach to decentralized fault detection*

167 K.A. GOPALA RAO, N. VISUANADHAM- *Sensor fault detection with reduced order models*

177 K.A. GOPALA RAO- *Control of a synchronous generator over wide range of operating conditions by eigenvalue placement*

187 K.A. GOPALA RAO- *Design of optimal discrete control systems with prescribed degree of stability*

193 K.A. GOPALA RAO- Decentralized fault detection in large-scale systems using secondary observers

207 C.M. ARORA, S.L. SURANA- Application of pattern recognition to power systems for transient security evaluation

219 M. de la SEN- Discrete quasi-ARMA modelling of systems with point delays

225 A. FUSARI- Public sector and economic growth. Estimation and simulation

237 K.J. SATAO- An information system for library management. A systems study and design

253 M. BUSH, D. PRESTON- A review of Japanese approaches to software production

265 S.R. VEDULA, Y. NARAHARI- Efficient scheduling of production and setups on a flexible machine using Lagrangian relaxation

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 3-13

ON CHARACTERISATION OF A DISTRIBUTION FUNCTION IN TERMS OF THE VALUES OF ITS CHARACTERISTIC FUNCTION AT SOME POINTS

BY

S. GHOSH

Department of Statistics, Presidency College,

Calcutta-700 073, India

1. ABSTRACT AND INTRODUCTION

A function $F : (-\infty, \infty) \to [0,1]$ is called a distribution function, if F is nondecreasing, right continuous and $\lim_{x \to -\infty} F(x) = 0 = \lim_{x \to \infty} (1 - F(x))$. The problem of characterisation of distribution functions is very important in Statistical theory. In this paper, some results in this direction are obtained. An implication of one of our findings in this paper is in the following direction. Suppose for each integer $j \in \{1, 2\}$, $F_j : (-\infty, \infty) \to [0,1]$ is a distribution function having characterestic function ϕ_j. We want to obtain a set of sufficient conditions under which, given a nonempty subset S of $(-\infty, \infty)$, $\phi_1(t) = \phi_2(t)$, for all $t \in S$, implies $F_1(x) = F_2(x)$, for all $x \in (-\infty, \infty)$. Some applications of our main results are also obtained. Our results seem to be new.

2. THE MAIN THEOREMS AND SOME OF THE APPLICATIONS.

A function $F : (-\infty, \infty) \to [0,1]$ is called a semidistribution function if F is nondecreasing, $\lim_{x \to -\infty} F(x) = \lim_{x \to \infty} (1 - F(x)) = 0$ and F is right continuous at 0. Suppose $g : (-\infty, \infty) \to (-\infty, \infty)$ is a continuous function. For all $a \in (-\infty, \infty)$, $b \in (a, \infty)$, $\int_{a \le x \le b} g \, dF$

will be interpreted as the Riemann-Stieltjes integral of g with respect to F on $[a, b]$ (cf. [1]). g is said to be generalised Riemann-Stieltjes integrable with respect to F on $(-\infty, \infty)$ if

$$\lim_{\substack{K_1 \to 0+,\ K_2 \to \infty \\ \text{and } K_1 < K_2}} \int_{K_1 \le x \le K_2} g \, dF \quad \text{as well as}$$

$$\lim_{K \to -\infty} \int_{K \le x \le 0} g \, dF$$

both exist finitely. In this case, the generalised Riemann-Stieltjes integral of g with respect to F will be defined as

$$\int_{-\infty}^{\infty} g \, dF = \lim_{\substack{K_1 \to 0+,\ K_2 \to \infty \\ \text{and } K_1 < K_2}} \int_{K_1 \le x \le K_2} g \, dF + \lim_{K \to -\infty} \int_{K \le x \le 0} g \, dF$$

Let $t \in (-\infty, \infty)$. Let $g_{1,t}(x) = \cos(tx)$, $g_{2,t}(x) = \sin(tx)$, for all $x \in (-\infty, \infty)$. Both $g_{1,t}$ and $g_{2,t}$ are generalised Riemann-Stieltjes integrable with respect to F. The semicharacterestic function of F at t will be defined as

$$\phi_F(t) = \int_{-\infty}^{\infty} \cos tx \, dF(x) + i \int_{-\infty}^{\infty} \sin tx \, dF(x), \quad (i = \sqrt{-1}).$$

Let $\gamma \in (0, \infty)$, $g_\gamma(x) = \exp(\gamma |x|)$, for all $x \in (-\infty, \infty)$. F is called regular at γ if the function g_γ is generalised Riemann-Stieltjes integrable with respect to F on $(-\infty, \infty)$.

With these notations, we formulate the following theorem.

THEOREM 1. Suppose for each integer $j \in \{1, 2\}$, $F_j : (-\infty, \infty) \to [0, 1]$ is a semi distribution function. We assume that,

(1) for all $x \in (-\infty, \infty)$, $(F_1 - F_2)$ is either right-continuous or left-continuous at x

(2) there exists $\gamma_0 \in (0, \infty)$ such that, for each integer $j \in \{1, 2\}$, F_j is regular at γ

(3) there exists a real number A which is a limit point of the set $\{t \in (-\infty, \infty) : \phi_1(t) = \phi_2(t)\}$

Then ,

$$F_1(x) = F_2(x)\ , \text{ for all } \quad x \in (-\infty, \infty)$$

Proof. (1) Let $j \in \{1, 2\}$. *We have that,* $\lim_{K \to \infty} (1-F_j(K))(\exp(\Upsilon_0 K))$ *= 0. By some known theorems (cf. [1], p.144 and 146),*

for all $K_1 \in (0, \infty)$, $K_2 \in (K_1, \infty)$,

$$\int_{K_1 \le x \le K_2} \exp(\Upsilon_0 x)\, d(1-F_j(x)) + \int_{K_1 \le x \le K_2} (1-F_j(x))\, \Upsilon_0 \exp(\Upsilon_0 x)\, dx$$

$$= (1-F_j(K_2)) \exp(\Upsilon_0 K_2) - (1-F_j(K_1)) \exp(\Upsilon_0 K_1)$$

One can easily conclude with the aid of monotone convergence theorem (cf.[5],p.124) that,

$$(1.1) \qquad \int_{0 < x < \infty} |F_1(x) - F_2(x)| (\exp(\Upsilon_0 x))\, dx < \infty$$

For all $j \in \{1, 2\}$, $\lim_{x \to -\infty} F_j(x)(\exp(\Upsilon_0 |x|)) = 0$. *Using this fact and proceeding in a very similar fashion, one can show that,*

$$(1.2) \qquad \int_{-\infty < x \le 0} |F_1(x) - F_2(x)| (\exp(\Upsilon_0 |x|))\, dx < \infty$$

Using equations (1.1) and (1.2), one can show that,

$$(1.3) \qquad \int_{-\infty < x < \infty} |F_1(x) - F_2(x)| (\exp(\Upsilon_0 |x|))\, dx < \infty$$

Let

$$(1.4) \qquad \Omega = (\alpha + i\beta \ : \ \text{where } \alpha \in (-\Upsilon_0, \Upsilon_0),\ \beta \in (-\infty, \infty)),$$

$$(1.5) \qquad F(z) = \int_{-\infty < x < \infty} (F_1(x) - F_2(x)) (\exp(z x))\, dx, \quad z \in \Omega$$

For all $\alpha \in (0, \infty)$ *and* $x \in (-\infty, \infty)$, $|x| \le \alpha^{-1} \exp(\alpha |x|)$

Using this fact and Canchy-Riemann equations (cf.[2], p.42), one can show that, F is an analytic function defined on Ω. *Using the hypothesis and the theory of Riemann-Stieltjes integration, one can show that,* A *is a limit point of the set*

$$\{t \in (-\infty, \infty) : \int_{-\infty < x < \infty} (F_1(x) - F_2(x)) (\exp(itx))\, dx = 0\}$$

By a known theorem (cf.[2],p.78), for all $\alpha \in (-\Upsilon_0, \Upsilon_0)$ *and* $\beta \in (-\infty, \infty)$, $F(\alpha + i\beta) = 0$. *Then for all* $\beta \in (-\infty, \infty)$,

$$(1.6) \qquad \int_{-\infty < x < \infty} (F_1(x) - F_2(x)) \exp(i\beta x)\, dx = 0$$

Using (1.3), one can easily show that,

(1.7) $$\int_{-\infty<x<\infty} |F_1(x) - F_2(x)| \, dx < \infty$$

Let m *denote Lebesgue-measure defined on Borel Subsets of* $(-\infty,\infty)$. *By a known theorem (cf. [7], p. 408),* $F_1 - F_2 = 0 \quad a.s. \; [m]$

Using the hypothesis of the theorem, one can show that,

$$F_1(x) = F_2(x), \text{ for all } x \in (-\infty,\infty)$$

The following result is known (cf.[8],p.212).

Suppose for each integer $j \in \{1,2\}$, $F_j : (-\infty,\infty) \to [0,1]$ *is a distribution function having characterestic function* ϕ_j . *We assume that, there exists* $r \in (0,\infty)$, $A \in (0,\infty)$ *such that, for all* $j \in \{1,2\}$,

$\int_{-\infty<x<\infty} \exp(r|x|) \, dF_j(x) < \infty$ *and,* $\phi_1(t) = \phi_2(t)$, *for all* $t \in (-A, A)$.

Then ,

$$F_1(x) = F_2(x), \text{ for all } x \in (-\infty,\infty)$$

The following result generalises the result stated above and is an immediate consequence of theorem 1.

COROLLARY 1.1. Suppose for each integer $j \in \{1,2\}$, $F_j : (-\infty,\infty) \to [0,1]$ *is a distribution function having characterestic function* ϕ_j . *We assume that, there exists* $r_0 \in (0,\infty)$ *such that, for each integer* $j \in \{1,2\}$,

$\int_{-\infty}^{\infty} \exp(r_0 |x|) \, dF_j(x) < \infty$ *and further, there exists a real number* A *whcih is a limit point of the set* $\{t \in (-\infty,\infty) : \phi_1(t) = \phi_2(t)\}$.

Then ,

$$F_1(x) = F_2(x), \text{ for all } x \in (-\infty,\infty)$$

To discuss our next theorem we need the following additional terminologies :

Here a, b *will denote real numbers such that,* $a < b$, *F will denote Borel-*σ*-algebra on* $[a,b]$, $\mu : F \to [0,\infty)$ *will denote a sigma-finite* measure. For all $t \in [1,\infty)$ denote, $L_t([a,b],F,\mu) = \{ f : f:[a,b]\to(-\infty,\infty)$ is a F-measurable function such that, $\int_{a\le x\le b} |f(x)|^t \mu(dx) < \infty\}$, $L_\infty([a,b],F,\mu) =$ $\{f : f:[a,b]\to(-\infty,\infty)$ is a F-measurable function such that, there exists $K(f)\in(0,\infty)$ such that, $|f|\le K(f) \text{ a.s.}[\mu]\}$. Let $t\in[1,\infty)$. For all $f\in L_t([a,b],F,\mu)$, denote $|f|_{t,\mu} = [\int_{a\le x\le b}|f(x)|^t\mu(dx)]^{t^{-1}}$. For all $g \in L_\infty([a,b],F,\mu)$, define $|g|_{\infty,\mu} = \text{Inf}\{K\in(0,\infty): |g|\le K \text{ a.s.}[\mu]\}$.

Here C_1 will denote theset of all complex numbers. For all $z\in C_1$, denote

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} \cdots + (-1)^n \frac{z^{2n}}{(2n)!} + \cdots ,$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} \cdots\cdots + (-1)^n \frac{z^{2n+1}}{(2n+1)!} + \cdots$$

(cf.[3], P.38). For the purpose of discussing our subsequent theorems, we need the following lemma which appears to be new.

LEMMA 1. Let $\mu' : F \to (-\infty,\infty)$ be a measure such that, $\mu'([a,b])=0$ and suppose for each $j \in \{1,2\}$ there is a sequence

$$(n_{j1}, n_{j2}, \cdots) \subset \{x \in (-\infty,\infty) : x \neq 0\} \text{ such}$$

that, $|n_{j1}| < |n_{j2}| < \cdots$, $\sum_{t=1}^{\infty} |n_{jt}|^{-1} = \infty$ and for each $t \in (1,2,\cdots)$,

$$\int_{a\le x\le b} f_j(n_{jt}x)\,\mu'(dx) = 0$$

where, for all $z \in C_1$, $f_1(z) = \sin z$, $f_2(z) = \cos z$

Then,

$$\mu'(A) = 0 \text{ , for all } A \in F$$

Proof.(1) Let $-\infty < a_1 < a$. Let F_1 denote Borel-σ-algebra on $[a_1, b]$. For all $A \in F_1$, denote $\mu_1(A) = \mu'(A \cap [a,b])$.

Then $\mu_1([a_1,b]) = 0$ *and* $\mu_1(\{a_1\}) = 0$. *For all* $x \in [a_1,b]$ *denote* $F(x) = \mu_1([a_1,x])$. *Then for all* $j \in \{1,2,\cdots\}$, *and* $t \in (1,2,\cdots)$,

$$\left|\int_{a_1 \le x \le b} \sin(n_{1t}x)\,F(dx)\right| + \left|\int_{a_1 \le x \le b} \cos(n_{2t}x)\,F(dx)\right| = 0,$$

where given a continuous function $f : [a_1,b] \to (-\infty,\infty)$,

$\int_{a_1 \le x \le b} f(x)\,F(dx)$ *should be interpreted as the Riemann Stieltjes* integral of f *with respect to F on* $[a_1,b]$ *(cf.[1]). For each integer* $j \in \{1,2\}$, *there exists a sequence* $(j_1 < j_2 < \cdots) \subset (1,2,\cdots)$ *such that,* $\sum_{t=1}^{\infty} (|n_{jj_t}|)^{-1} = \infty$, *and either, for each* $t \in (1,2,\cdots)$, $n_{jj_t} > 0$ *or for each* $t \in (1,2,\cdots)$, $n_{jj_t} < 0$. *Further for all* x, $\cos(-x) = \cos x$, $\sin(-x) = -\sin x$. *Using this fact, one can assume without loss of generality that, for each* $t \in (1,2,\cdots)$, $n_{1t} > 0$, $n_{2t} > 0$. *Using the fact that,* $F(a_1) = F(b) = 0$ *and some theorems involving Riemann-Stieltjes integration (cf. [1], p.144 and 146), one can show that, for each* $t \in (1,2,\cdots)$,

$$\left|\int_{a_1 \le x \le b} F(x)\cos(n_{1t}x)\,dx\right| + \left|\int_{a_1 \le x \le b} F(x)\sin(n_{2t}x)\,dx\right| = 0. \tag{1}$$

Let $\Omega_1 = [z : z = \alpha + i\beta$, *where* $\alpha \in (-1,\infty)$, $\beta \in (0,\infty)]$, *where* $i = \sqrt{-1}$. *Let* $\Omega_2 = [z : z = \alpha + i\beta$, *where* $\alpha \in (-1,\infty)$, *and* $\beta = 0]$. *Let* $\Omega_3 = [z : z = \alpha + i\beta$, *where* $\alpha \in (-1,\infty)$ *and* $\beta \in (-\infty,\infty)]$. *For all* $z \in \Omega_1 \cup \Omega_2$ *and for all* $j \in \{1,2\}$, *define*

$$F_j(z) = \left[\exp(-iz(|b| + |a_1|))\right]^{-1} \int_{a_1 \le x \le b} F(x)\,f_j(zx)\,dx. \tag{2}$$

Let $z = \alpha + i\beta$, *where* $\alpha \in (-1,\infty)$, $\beta \in (0,\infty)$. *Let* $x \in [a_1,b]$.

By a known result (cf. [2], p. 31),

$$|\sin((\alpha+i\beta)x)| = |\sin(\alpha x)\cos(i\beta x) + \cos(\alpha x)\sin(i\beta x)|$$

$$\leq 2[\exp(\beta(|b|+|a_1|))] = 2|\exp(-i(|b|+|a_1|)z)| ,$$

$$|\cos((\alpha+i\beta)x)| = |\cos(\alpha x)\cos(i\beta x) - \sin(\alpha x)\sin(i\beta x)|$$

$$\leq 2|\exp(-i(|b|+|a_1|)z)| .$$

Let $j \in \{1, 2\}$. *Restriction of* F_j *to* Ω_1 *is a bounded analytic function. Further if* $\{z_1, z_2, \ldots\ldots\} \subset \Omega_1$ *be such that,* $z_0 = \lim_{n\to\infty} z_n$ *exists and is an element of* Ω_2, $F_j(z_0) = \lim_{n\to\infty} F_j(z_n)$ *and further,* $F_j(z_0)$ *is real. By the Schwarz reflection principle. (cf. [10] p.260), there is a bounded analytic function* $F_j' : \Omega_3 \to C_1$ *such that,* $F_j'(z) = F_j(z)$, *for all* $z \in \Omega_1 \cup \Omega_2$.

Let $\Omega_4 = [z : z = \alpha + i\beta$, *where* α, β *are both real and* $(\alpha^2+\beta^2) < 1]$. *Define* $g_j(z) = F_j'\left(\frac{1+z}{1-z}\right)$, *for all* $z \in \Omega_4$. g_j *is a bounded analytic function defined on* Ω_4, *and for all* $t \in (1, 2, \ldots\ldots)$, $g_j\left(\frac{n_{jt}-1}{n_{jt}+1}\right) = 0$

$\sum_{t=1}^{\infty}\left[1 - \left|\frac{n_{jt}-1}{n_{jt}+1}\right|\right] = \infty$. *Hence by a known theorem (cf. [10],p.335),*

$g_j(z) = 0$, *for all* $z \in \Omega_4$.

So,

(1.3) $F_j'(z) = 0$, *for all* $z \in \Omega_3$

(2) For all $A \in F_1$, *define* $\mu_1^+(A) = \operatorname{Sup}\{\mu_1(B) : B \subset A$ *and* $B \in F_1\}$, $\mu_1^-(A) = \operatorname{Sup}\{-\mu_1(B) : B \subset A$ *and* $B \in F_1\}$.

By the hypothesis of the theorem and a known result (cf.[9], p.104),

$\mu_1^+([a,b]) = \mu_1^-([a,b]) = \alpha$ *(say). One can assume without loss of generality that,* $\alpha \in (0,\infty)$, *because if* $\alpha = 0$, *by a known theorem (cf.[9],p.104),* $\mu_1(A) = 0$, *for all* $A \in F_1$ *and so,* $\mu'(A) = 0$, *for all* $A \in F$. *By equation (1.3), for all* $t \in [1,2]$,

(2.1) $$\alpha^{-1} \int_{a_1 \le x \le b} (\exp(itx))\, \mu_1^+(dx) = \alpha^{-1} \int_{a_1 \le x \le b} (\exp(itx))\, \mu_1^-(dx).$$

By theorem 1 and a known result (cf. [9], p.104),

$\mu_1(A) = 0$, *for all* $A \in F_1$ *and so,* $\mu'(A) = 0$, *for all* $A \in F$.

The following result appears to be new.

THEOREM 2. Let $f : [a,b] \to (-\infty,\infty)$ *be a continuous function, let* $p \in [1,\infty)$ *and let* $h \in L_p([a,b], F, \mu)$ *be such that,* $\mu(x \in [a,b] : h(x) = 0) = 0$. *Let* $g \in L_p([a,b], F, \mu)$.
Suppose for each integer $j \in \{1,2\}$, $(n_{j1}, n_{j2}, \cdots\cdots)$ *is a sequence of nonzero real numbers such that* $|n_{j1}| < |n_{j2}| < \cdots$ *and* $\sum_{t=1}^{\infty} |n_{jt}|^{-1} = \infty$.
Then given $\delta \in (0,\infty)$, *there are integers* $n_1(f,\delta) \ge 1$, $n_2(g,\delta) \ge 1$ *and*
$\{a(0,f,\delta), b(0,g,\delta), a(1,j,f,\delta)\ a(2,j,f,\delta), j=1,\cdots n_1(f,\delta), b(1,j,g,\delta),$

$b(2,j,g,\delta), j=1, \cdots\cdots\cdots n_2(g,\delta)\} \subset (-\infty,\infty)$ *such that,*

(1) $$\sup_{x \in [a,b]} \left| f(x) - a(0,f,\delta) - \sum_{j=1}^{n_1(f,\delta)} [a(1,j,f,\delta)\cos(n_{1j}x) + a(2,j,f,\delta)\sin(n_{2j}x)] \right| \le \delta,$$

(2) $$\int_{a \le x \le b} \left| g(x) - \{b(0,g,\delta) + \sum_{j=1}^{n_2(g,\delta)} [b(1,j,g,\delta)\cos(n_{1j}x) + b(2,j,g,\delta)\sin(n_{2j}x)]\} h(x) \right|^p \mu(dx)$$
$$\le \delta$$

Proof. (1) Let $C([a,b])$ denote the set of all real-valued continuous functions defined on $[a,b]$. For all $f_1 \in C([a,b])$, denote

$\|f\|_1 = \sup_{x \in [a,b]} |f_1(x)|$. $C([a,b])$ is a normed real linear space. If equation (1) in the statement of theorem 2 is false, by the Hahn Banach theorem (cf. [11], p.229), some integral representation theorems of continuous linear functionals on $C([a,b])$ (cf.[6],p.249 and [10],p.42) and a known argument (cf. [5]), there exists a measure $\mu' : F \to (-\infty, \infty)$ and a function $f \in C([a,b])$ such that, $\mu'([a,b]) = 0$,

$\int_{a \le x \le b} f(x)\,\mu'(dx) = 1$ but for each integer $t = 1, 2, \cdots$,

(1.1) $\left| \int_{a \le x \le b} \cos(n_{1t}\, x)\,\mu'(dx) \right| + \left| \int_{a \le x \le b} \sin(n_{2t}\, x)\,\mu'(dx) \right| = 0$

By lemma 1, $\mu'(A) = 0$, for all $A \in F$ and so,

$\int_{a \le x \le b} f(x)\mu'(dx) = 0$. This contradiction proves the truth of equation (1) in the statement of theorem 2.

(2) If $P \in (1, \infty)$, define Q by the equation $P^{-1} + Q^{-1}$ = 1. If P = 1, Q should be interpreted to be ∞. Suppose equation (2) in the statement of theorem 2 is false. Proceeding as in the proof of theorem 3.2 in([4], p.26), one can easily show that, there exists $f \in L_P([a,b], F, \mu)$, $g \in L_Q([a,b], F, \mu)$ such that,

(2.1) $$\int_{a \le x \le b} f(x)\, g(x)\, \mu(dx) = 1$$

but, for each $t \in (1, 2, \cdots)$, $j \in \{1, 2\}$,

(2.2) $\left| \int_{a \le x \le b} h(x)\, g(x)\, \mu(dx) \right| + \left| \int_{a \le x \le b} h(x)\, g(x)\, f_j(n_{jt}\, x)\, \mu(dx) \right|$

= 0, where for all $x \in (-\infty, \infty)$, $f_1(x) = \cos x$, $f_2(x) = \sin x$

For each $A \in F$, define $\mu'(A) = \int_A h(x)\, g(x)\, \mu(dx)$

Equation (2.2), the fact that, $\mu'([a,b]) = 0$, and lemma 1 together implies that, $\mu'(A) = 0$, for all $A \in F$. The hypothesis of the theorem implies that, $\mu(x \in [a,b] : g(x) \neq 0) = 0$. So,

$$\int_{a \le x \le b} f(x)\, g(x)\, \mu(dx) = 0 \tag{2.3}$$

Equations (2.1) and (2.3) contradict one another. Thus equation (2) in the statement of theorem 2 holds.

The following results easily follows from lemma 1.

COROLLARY 2. Suppose for each integer $j \in \{1, 2\}$, $F_j : [a,b] \to [0,1]$ is a distribution function such that, $F_j(a) = 0 = 1 - F_j(b)$. Let

$(n_1 < n_2 < \cdots\cdots) \subset (0, \infty)$ be such that,

$\sum_{t=1}^{\infty} (n_t)^{-1} = \infty$ and for all $t \in (1, 2, \cdots\cdot)$,

$$\int_{a \le x \le b} (\exp(i n_t x))\, dF_1(x) = \int_{a \le x \le b} (\exp(i n_t x))\, dF_2(x).$$

Then,

$$F_1(x) = F_2(x), \text{ for all } x \in [a,b]$$

REFERENCES

1. APOSTOL, T. M. (1973) : "Mathematical Analysis", Second Edition, Narosa Publishing House, New DELHI.

2. CHATTERJEE, P.C. (1964) : "Higher Algebra", Bookland Private Limited, Calcutta.

3. CONWAY, JOHN, B. (1980) : "Functions of one complex variable", Second Edition, Springer Verlag, New-York.

4. GHOSH, S. (1991) : "Some results related to characterisation of distribution functions and some applications", Proceedings of the international A.M.S.E. Conference held in New-Delhi, India during December 9-11, 1991. Volume-I, P.19-31, A.M.S.E. Press, France.

5. GHOSH, S. : "A probabilistic approach to Stone-Weierstrass approximation theorem", submitted for publication in the international A.M.S.E. Conference to be held in Bangalore, India during December 8-10, 1993.

6. HALMOS, P.R. (1964) : "Measure theory", Van Nostrand, Princeton.

7. HEWITT, EDWIN AND STROMBERG, KARL (1978) : "Real and abstract analysis", Narosa Publishing House, New Delhi.

8. LOEVE, M. (1963) : "Probability theory", third edition, Van Nostrand,Princeton.

9. NEVEU, J. (1965) : "Mathematical Foundations of the Calculus of Probability", Holden Day, San-Fransisco.

10. RUDIN, W. (1974) : "Real and complex analysis", Second edition, Tata McGraw Hill Publishing Company, New Delhi.

11. SIMMONS, G.F. (1963) : "Introduction to Topology and modern analysis", McGraw Hill, New York.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 15-24

SHARPENING OF A THEOREM OF GWO DONG LIN AND AN APPLICATION

BY

S. GHOSH

Department of Statistics, Presidency College,
Calcutta-700 073, India

1. *ABSTRACT AND INTRODUCTION*

Let (Ω, σ, P) *be a probability-space. Let* $X : \Omega \to (-\infty, \infty)$ *be a* σ*-measurable function. Let* $F(x) = P(\omega \in \Omega : X(\omega) \leq x)$, *for all* $x \in (-\infty, \infty)$. *F is called the distribution function of X. It is an important problem in Statistics to identify a distribution function in terms of some of its moments. In this paper, we have obtained a slight sharpening of a result of GWO Dong Lin (1992) in this direction. An application of our main result is also obtained. Our results appear to be new. Let* $Y : \Omega \to [0, \infty)$, $Y_1 : \Omega \to [0, \infty)$ *be* σ*-measurable functions. If for some* $t \in (-\infty, 0)$, $E(Y^t) < \infty$, $E(|\log_e Y_1|) < \infty$, *then it is automatically understood that,*

$$P(\omega \in \Omega : Y(\omega) = 0) + P(\omega \in \Omega : Y_1(\omega) = 0) = 0$$

2. *THE MAIN THEOREM AND AN APPLICATION.*

THEOREM 1. Let (Ω, σ, P) *be a probability-space. Suppose for each* $i \in \{1, 2\}$, $X_i : \Omega \to [0, \infty)$ *is a* σ*-measurable function and suppose* F_i *is the distribution function of* X_i . *We assume that, there exists* $\alpha_1 \in (-\infty, \infty)$, $\alpha_2 \in (\alpha_1, \infty)$,
such that,

(i) for all $i \in \{1,2\}$ and for all $\beta \in (\alpha_1, \alpha_2)$, $E(X_i^{\beta}) < \infty$

(ii) there exists a sequence $\{\beta_0, \beta_1, \beta_2, \cdots\}$ of nonZero distinct real numbers such that, each $\beta_i \in (\alpha_1, \alpha_2)$, $\lim_{n \to \infty} \beta_n = \beta_0$ and further, for each $n \in (1, 2, \cdots)$, $E(X_1^{\beta_n}) = E(X_2^{\beta_n})$

Then,

$$F_1(x) = F_2(x) \quad , \quad \text{for all} \quad x \in (-\infty, \infty)$$

On the other hand, if $\alpha_1 = 0$, $\alpha_2 = 1$ and for each $i \in \{1, 2\}$, $E(|X_i| + |\log_e X_i|) < \infty$, the assumption that, "$\beta_0 \in (0,1)$" can be replaced by the assumption that, "$\beta_0 \in (0,1]$".

Proof. (1) Let $x \in [0, \infty)$. Define $G_1(x) = \int_{0 \le t \le x} t^{\beta_1} F_1(dt)$, $G_2(x) = \int_{0 \le t \le x} t^{\beta_1} F_2(dt)$, $G_1(\infty) = \lim_{y \to \infty} G_1(y)$, $G_2(\infty) = \lim_{y \to \infty} G_2(y)$. Let $\alpha \in (\alpha_1 - \beta_1, \alpha_2 - \beta_1)$ be such that, $\alpha \neq 0$. If $\alpha < 0$, $\limsup_{t \to 0+} [t^{\alpha} G_1(t)] \le \limsup_{t \to 0+} [\int_{0 \le x \le t} x^{\alpha + \beta_1} F_1(dx)] = 0$. If $\alpha > 0$, by virtue of the fact that, $\lim_{t \to 0+} [\int_{0 \le x \le t} x^{\beta_1} F_1(dx)] = 0$, $\limsup_{t \to 0+} [t^{\alpha} G_1(t)] = 0$. Similarly, $\limsup_{t \to 0+} [t^{\alpha} G_2(t)] = 0$. If $\alpha > 0$, $\limsup_{t \to \infty} [t^{\alpha} (G_1(\infty) - G_1(t))] \le \limsup_{t \to \infty} [\int_{t < x < \infty} x^{\alpha + \beta_1} F_1(dx)] = 0$. If $\alpha \in (-\infty, 0)$, $\limsup_{t \to \infty} [t^{\alpha} (G_1(\infty) - G_1(t))]$

$$\le \limsup_{t \to \infty} [\int_{t < x < \infty} x^{\beta_1} F_1(dx)] = 0.$$

Thus for all $i \in \{1,2\}$,

$$\limsup_{t \to \infty} [t^{\alpha} (G_i(\infty) - G_i(t))] = 0 .$$

Thus for all $i \in \{1,2\}$, and for each nonzero real number $\alpha \in (\alpha_1 - \beta_1, \alpha_2 - \beta_1)$,

$$\text{(1.1)} \quad \limsup_{t \to 0+} [t^{\alpha} G_i(t)] = \limsup_{t \to \infty} [t^{\alpha} (G_i(\infty) - G_i(t))] = 0$$

Let $\alpha \in (\alpha_1-\beta_1,\ \alpha_2-\beta_1)$ *be such that,* $\alpha \neq 0$. *Let* $t \in (0,1)$, $i \in \{1,2\}$. *Using the theory of Riemann-Stieltjes integration (cf.[1], P.144 and 146), one can show that,*

$$\int_{t\le x\le 1} x^{\alpha}\, dG_i(x) + \int_{t\le x\le 1} G_i(x)\,\alpha\, x^{\alpha-1}\, dx = G_i(1) - t^{\alpha} G_i(t)$$

Using (1.1) and monotone convergence theorem (cf.[6],p.41), one can easily show that,

$$\int_{0\le x\le 1} G_i(x)\, x^{\alpha-1}\, dx < \infty \tag{1.2}$$

Thus,

$$\int_{0\le x\le 1} |G_1(x) - G_2(x)|\, x^{\alpha-1}\, dx < \infty \tag{1.3}$$

Let $t \in (1,\infty)$, $i \in \{1,2\}$. *Let* $g(x) = x^{\alpha}$, *for all* $x \in [1,\infty)$.

By the theory of Riemann-Stieltjes integration (cf.[1], P.144 and 146),

$$\int_{1\le x\le t} g\, d(G_i(\infty) - G_i) + \int_{1\le x\le t} (G_i(\infty) - G_i(x))\,\alpha\, x^{\alpha-1}\, dx$$

$$= t^{\alpha}(G_i(\infty) - G_i(t)) - (G_i(\infty) - G_i(t))$$

where, $\int_{1\le x\le t} g\, d(G_i(\infty) - G_i)$ *is to be interpreted as the Riemann-Stieltjes integral of* g *with respect to the function* $(G_i(\infty) - G_i)$ *on [1,t]. Using this, (1.1) and monotone convergence theorem, one can easily show that, for all* $i \in \{1,2\}$, $\int_{1\le x<\infty} (G_i(\infty) - G_i(x))\, x^{\alpha-1}\, dx < \infty$

Using this and equation (1.3), one can easily show that,

for each nonzero $\alpha \in (\alpha_1-\beta_1,\ \alpha_2-\beta_1)$,

$$\int_{0\le x<\infty} x^{\alpha-1}\, |G_1(x) - G_2(x)|\, dx < \infty \tag{1.4}$$

Let $i \in \{1,2\}$. *Let* $\alpha < \beta_1' < \beta_1 < \beta_1'' < \alpha_2$, *Where if* $\beta_1 > 0$, $\beta_1' > 0$. $\int_{0<t\le 1} (-\log_e t)\, G_i(dt) = \int_{0<t\le 1} (-\log_e t)\, t^{\beta_1}\, F_i(dt)$

$$= \int_{0<t\le 1} (-\log_e t)\, t^{\beta_1-\beta_1'}\, t^{\beta_1'}\, F_i(dt)$$

There exists $k_1 \in (0,\infty)$ *such that,* $\sup_{0<t\le 1} \left|(-\log_e t)\, t^{\beta_1-\beta_1'}\right| \le k_1$ *Then*

(1.5) $\int_{0<t\le 1} (-\log_e t)\, G_i(dt) \le K_1 \int_{0\le t\le 1} t^{\beta_i'}\, F_i(dt) < \infty$

Thus, $\lim_{t\to 0+} \left[|\log_e t|\, G_i(t) \right] = 0$

$$\int_{1\le t<\infty} (\log_e t)\, G_i(dt) = \int_{1\le t<\infty} (\log_e t)\, t^{\beta_1}\, F_i(dt)$$

$= \int_{1\le t<\infty} (\log_e t)\, t^{\beta_1-\beta_1''}\, t^{\beta_1''}\, F_i(dt)$. There exists $K_2\in(0,\infty)$ such that, $\sup_{1\le t<\infty} \left| (\log_e t)\, t^{\beta_1-\beta_1''} \right| \le K_2$. Then,

(1.6) $\int_{1\le t<\infty} (\log_e t)\, G_i(dt) \le K_2 \int_{1\le t<\infty} t^{\beta_i''}\, F_i(dt) < \infty$

So for all $i\in\{1,2\}$, $\lim_{t\to\infty} \left[(G_i(\infty)-G_i(t)) \log_e t \right] = 0$. Thus for all $i \in \{1,2\}$,

(1.7) $\lim_{t\to 0+} \left[|\log_e t|\, G_i(t) \right] = \lim_{t\to\infty} \left[\log_e t\, (G_i(\infty)-G_i(t)) \right] = 0.$

Let $i\in\{1,2\}$. Let $m\in(0,1)$. Using the theory of Riemann Stieltjes integrals (cf.[1], p.144 and 146), one can easily show that,

$\int_{m\le t\le 1} (-\log_e t)\, G_i(dt) + \int_{m\le t\le 1} G_i(t)(-t^{-1})\, dt = G_i(m) \log_e m$ and by (1.5), $G_i(m)\, |\log_e m| \to 0$ as $m\to\infty$. Using this, monotone convergence theorem and equation (1.5), one can show that,

(1.8) $$\int_{0<t\le 1} G_i(t)\, t^{-1}\, dt < \infty .$$

Let $g_2(t) = \log_e t$, for all $1\le t<\infty$. Using equation (1.7) and the theory of Riemann-Stieltjes integration (cf. [1], p.144 and 146), one can easily show that,

$$\lim_{K\to\infty} \left[\int_{1\le t\le K} g_2\, d(G_i(\infty)-G_i) + \int_{1\le t<\infty} (G_i(\infty)-G_i(t))\, t^{-1}\, dt \right]$$

$= 0$, where $\int_{1\le t\le K} g_2\, d(G_i(\infty)-G_i)$ should be interpreted as the Riemann-Stieltjes of g_2 with respect to $(G_i(\infty)-G_i)$ on [1,K]. Using the above fact, equation (1.6) and monotone convergence theorem, one can easily show that, (1.9) $$\int_{1\le t<\infty} (G_i(\infty) - G_i(t))\, t^{-1}\, dt < \infty$$

Using equations (1.4), (1.8) and (1.9), one can easily show that, for each

$$\alpha \in (\alpha_1 - \beta_1, \alpha_2 - \beta_1),$$

(1.10) $$\int_{0\le x<\infty} x^{\alpha-1} |G_1(x) - G_2(x)|\, dx < \infty$$

Let $\alpha \in (\alpha_1 - \beta_1, \alpha_2 - \beta_1)$. Let $\alpha' \in (\alpha_1 - \beta_1, \alpha)$, $\alpha'' \in (\alpha, \alpha_2 - \beta_1)$.

There exists $K \in (0,\infty)$, $K_1 \in (0,\infty)$ such that,

$$\sup_{0<x\le 1} \left| x^{\alpha-\alpha'} (\log_e x) \right| \le K, \quad \sup_{1\le x<\infty} \left| x^{\alpha-\alpha''} \log_e x \right| \le K_1$$ Then,

$$\int_{0<x<\infty} x^{\alpha-1} |G_1(x) - G_2(x)|\, |\log_e x|\, dx$$
$$\le K \int_{0<x\le 1} x^{\alpha'-1} |G_1(x) - G_2(x)|\, dx + K_1 \int_{1\le x<\infty} x^{\alpha''-1} |G_1(x) - G_2(x)|\, dx$$
$<\infty$. Thus, for each $\alpha \in (\alpha_1 - \beta_1, \alpha_2 - \beta_1)$,

(1.11) $$\int_{0<x<\infty} x^{\alpha-1} |G_1(x) - G_2(x)|\, |\log_e x|\, dx < \infty$$

(2) Let $i \in (2, 3, \cdots)$, $\alpha_i = \beta_i - \beta_1$. Let $g_1(x) = x^{\alpha_i}$ for all $x \in [0,\infty)$, where if $\alpha_i < 0$, $g_1(0)$ should be interpreted to be 0.

$$\int_{0\le x<\infty} x^{\beta_i} F_1(dx) = \int_{0\le x<\infty} g_1(x)\, G_1(dx) = \int_{0\le x<\infty} g_1(x)\, G_2(dx) < \infty.$$ So,

$$\lim_{\substack{K_1\to 0+,\, K_2\to\infty \\ \text{and } K_1 < K_2}} \int_{K_1\le x\le K_2} g_1\, d(G_1 - G_2) = 0$$

Where $\int_{K_1\le x\le K_2} g_1\, d(G_1 - G_2)$ should be interpreted as the Riemann Stieltjes integral of g_1 with respect to $(G_1 - G_2)$ on $[K_1, K_2]$. Let $0 < K_1 < K_2 < \infty$.

By ([1], p.144 and 146), $\int_{K_1\le x\le K_2} g_1\, d(G_1 - G_2) + \int_{K_1\le x\le K_2} (G_1(x) - G_2(x))\, \alpha_i x^{\alpha_i - 1}\, dx$

$$= K_2^{\alpha_i} (G_1(K_2) - G_2(K_2)) - K_1^{\alpha_i} (G_1(K_1) - G_2(K_2))$$

$\longrightarrow 0$ as $K_1 \longrightarrow 0+$, $K_2 \longrightarrow \infty$ (cf. equation (1.1))

Using the above fact, equation (1.1) and monotone convergence theorem, one can easily show that, for all $i \in (2,3,\ldots)$,

(2.1) $$\int_{0<x<\infty} x^{\alpha_i - 1}\,(G_1(x) - G_2(x))\,dx = 0$$

(3) Let $i = \sqrt{-1}$ and

(3.1) $$\Omega_1 = [z : z = \alpha + i\beta, \text{ where } \alpha \in (\alpha_1 - \beta_1, \alpha_2 - \beta_1), \beta \in (-\infty, \infty)],$$

(3.2) $$F(z) = \int_{0<x<\infty} (F_1(x) - F_2(x))\,x^{z-1}\,dx, \qquad z \in \Omega_1$$

Using equations (1.10), (1.11) and the Cauchy-Riemann equations (cf. [2], p 42) one can easily show that, F is an analytic function defined on the region Ω_1. By a known theorem (cf. [2], p.78), $F(z) = 0$, for all $z \in \Omega_1$, and so, for all $\beta \in (-\infty, \infty)$,

(3.3) $$\int_{0<x<\infty} (G_1(x) - G_2(x))\,x^{-1}\,[\exp(i\beta \log_e x)]\,dx = 0$$

Using the theory of Riemann-Stieltjes integrals, dominated convergence theorem (cf. [6], p.42) and the fact that,

$\int_{0<x<\infty} |G_1(x) - G_2(x)|\,x^{-1}\,dx < \infty$ (cf. equation (1.11)), one can easily show that, for all $\beta \in (-\infty, \infty)$,

(3.4) $$\int_{0<x<\infty} (\exp[i\beta \log_e x])\,G_1(dx) = \int_{0<x<\infty} (\exp[i\beta \log_e x])\,G_2(dx).$$

By a known theorem (cf. [5], p. 186),

(3.5) $$G_1(x) = G_2(x), \quad \text{for all } x \in [0, \infty) .$$

(4) Let σ_1 = Borel-σ-algebra on $[0,\infty)$. For each $i \in \{1,2\}$, denote by m_i, the unique probability-measure on $([0,\infty), \sigma_1)$ such that, for all $x \in (0,\infty)$, $F_i(x) = m_i([0, x])$ (cf. [5], p.167). Let $m(A) = 2^{-1}(m_1(A) + m_2(A))$, for all $A \in \sigma_1$. For each $i \in \{1,2\}$,

let $f_i : [0,\infty) \to [0,\infty)$ be a σ_1-measurable function such that, for all $A \in \sigma_1$,

$$m_i(A) = \int_A f_i(x)\, m(dx)$$ (cf.[6], p. 108).

One can easily show that, for all $x \in (0,\infty)$,

$$G_1(x) = \int_{0 \le t \le x} t^{\beta_1} F_1(dt) = \int_{0 \le t \le x} t^{\beta_1} f_1(t)\, m(dt)$$

$$= G_2(x) = \int_{0 \le t \le x} t^{\beta_1} f_2(t)\, m(dt).$$

Using this, one can easily show that, there exists $N \in \sigma_1$ such that, $m(N) = 0$ and for each nonzero $x \in N^c$, $f_1(x) = f_2(x)$. Let $x \in (0,\infty)$.

$$m_1((0,x]) = \int_{0 < t \le x} f_1(t)\, m(dt) = \int_{0 < t \le x} f_2(t)\, m(dt)$$

$$= m_2((0,x]).$$ Since $m_1((0,\infty)) = m_2((0,\infty))$ and $1 = m_1([0,\infty)) = m_2([0,\infty))$, $m_1(\{0\}) = m_2(\{0\})$.

Thus for all $x \in [0,\infty)$, $F_1(x) = F_2(x)$.

(5) Assume now, $\alpha_1 = 0$, $\alpha_2 = \beta_0 = 1$ and for all $i \in \{1,2\}$, $E(|X_i| + |\log_e X_i|) < \infty$. Let $i \in \{1,2\}$. Since $\int_{0 < x < 1} (-\log_e x) F_i(dx) < \infty$, $\lim_{t \to 0+} [|\log_e t| F_i(t)] = 0$. Since $\int_{1 \le t < \infty} (\log_e t) F_i(dt) < \infty$, $\lim_{t \to \infty} [\log_e t (1 - F_i(t))] = 0$. Thus for each $i \in \{1,2\}$,

(5.1) $$\lim_{t \to 0+} [|\log_e t| F_i(t)] = 0 = \lim_{t \to \infty} [\log_e t (1 - F_i(t))].$$

Using these facts and the theory of Riemann-Stieltjes integration (cf.[1],p.144 and 146), one can show that, for each $i \in \{1,2\}$,

$$\int_{0 < x \le 1} x^{-1} F_i(dx) + \int_{1 \le x < \infty} x^{-1} (1 - F_i(x))\, dx < \infty,$$ and thus,

(5.2) $$\int_{0 < x < 1} x^{-1} |F_1(x) - F_2(x)|\, dx < \infty$$

Since for each $i \in \{1, 2\}$, $\int_{0<x<\infty} x \, F_i(dx) < \infty$, *using the theory of Riemann-Stieltjes integration, one can easily show that,*

(5.3) $$\int_{0<x<\infty} |F_1(x) - F_2(x)| \, dx < \infty$$

Let

(5.4) $\Omega_2 = [\, Z : Z = \alpha + i\beta$, *where* $\alpha \in (-\infty, \infty)$, $\beta \in (-\infty, \infty)$ *and* $(\alpha^2 + \beta^2)^{1/2} < 1 \,]$,

(5.5) $$F_1'(Z) = \int_{0<x<\infty} (F_1(x) - F_2(x)) \, x^{Z-1} \, dx , \; Z \in \Omega_2$$

Using (5.2) and (5.5), one can easily show that, F_1' *is an analytic uniformly bounded function defined on* Ω_2. *By the hypothesis of the theorem, there exists a sequence* $\{\beta_1, \beta_2, \cdots\}$ *of distinct strictly positive real numbers such that, each* $\beta_i < 1$

and,

(5.6) $$\int_{0<x<\infty} (F_1(x) - F_2(x)) \, x^{\beta_i - 1} \, dx = 0$$

$\sum_{i=1}^{\infty} (1 - (1 - \beta_i)) = \infty$ *and so by a known theorem (cf.[7], p.305),* $F_1'(Z) = 0$, *for all* $Z \in \Omega_2$. *So* , $E(X_1^{\beta}) = E(X_2^{\beta})$, *for all* $\beta \in [3^{-1}, 2^{-1}]$ *and so, by what has already been proved,*

$$F_1(x) = F_2(x), \quad \text{for all} \quad x \in (-\infty, \infty)$$

GWO Dong Lin (cf.[4], p.128) has proved the following theorem.

Let (Ω, σ, P) *be a probability-space. Let* $Y : \Omega \to (0, \infty)$ *be a* σ*-measurable function such that, for some* $\alpha \in (0, \infty)$, $\int Y^{\alpha} \, dP < \infty$. *Let* $\{\alpha_1, \alpha_2, \cdots\cdots\} \subset (0, \alpha)$ *be a sequence of distinct real numbers converging to some point* $\alpha_0 \in (0, \alpha)$. *Then the sequence of moments* $\{E(Y^{\alpha_m}), m = 1, 2, \cdots\}$ *characterises the distribution function of Y.*

Clearly theorem 1 is a sharpening of the above result.

The following result can be easily proved using theorem 1.

COROLLARY 1. Let (Ω, σ, P) be a probability-space. Suppose for each integer $i \in \{1,2\}$, $X_i : \Omega \to (0,\infty)$ is a σ-measurable function. Suppose I is an open interval such that, $0 \in I$ and for all $\alpha \in I$ and for each $i \in \{1,2\}$, $E(X_i^{\alpha}) < \infty$. We assume that, there exists a bounded sequence $\{\alpha_1, \alpha_2, \cdots\}$ of real numbers such that, for each integer $j \geq 1$, $\alpha_j \neq 0$ and

$$E\left[\exp(i\alpha_j \log_e X_1)\right] = E\left[\exp(i\alpha_j \cdot \log_e X_2)\right]$$

Then X_1 and X_2 are identically distributed.

The argument that has to be used for proving Corollary 1 is similar to the one used in the proof of theorem 1 and hence the proof of Corollary 1 is omitted. In this connection it should be noted that, an implication of the assumption that $0 \in I$ is that, for all $i \in \{1,2\}$, $E(|\log_e X_i|) < \infty$ and consequently, for all $\alpha \in I$,

$$\int_{0 < x < \infty} |F_1(x) - F_2(x)| \, x^{\alpha - 1} (1 + |\log_e x|) \, dx < \infty$$

R E F E R E N C E S

1. APOSTOL, T.M. (1973) : "Mathematical analysis", Second edition, Narosa Publishing House, New Delhi.

2. CONWAY, JOHN B. (1980) : "Functions of one complex variable", Second edition, Springer Verlag, New York.

3. GHOSH, S. (1992) : "Some results related to characterisation of distribution functions and some applications", Proceedings of the 1991 international A.M.S.E. Conference held in New Delhi, India during December 9-11, 1991, Volume I, p.19-31, A.M.S.E Press, France.

4. LIN, GWO. DON.(1992) : "Characterisations of distributions via moments", Sankhya, Volume 54, Series A, Part I, 128-132.

5. LoEVE, M.(1963) : "Probability theory", third edition, Van Nostrand, Princeton.

6. NEVEU, J. (1965) : "Mathematical Foundations of the Calculus of Probability", Holden Day, San-Fransisco.

7. RUDIN, W. (1974) : "Real and Complex analysis", Second edition, McGraw Hill, New-YOrk.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 25-37

A SET OF NECESSARY AND SUFFICIENT CONDITIONS FOR PERFECT MIXTURE OF PERFECT MEASURES BEING PERFECT

By

S. GHOSH

Department of Statistics, Presidency College,

Calcutta - 700 073, India

ABSTRACT

In this paper, an analogue of the concept of perfectness for Banach-Space-Valued measures has been dealt with and a necessary and sufficient condition for perfect mixture of perfect transition probabilities being perfect has been developed. A theorem of Ramachandran (1973) has been generalised with the aid of our theorem.

1. INTRODUCTION

It is known that, perfect mixture of discrete transition probabilities is perfect, but perfect mixture of perfect transition-probabilities need not be perfect (cf.[6]). Suppose (E,F,P) is a perfect probability-space, (Ω_2, σ_2) *is a measurable space and H is a real Banach-Space norm on which is donated by* $\|\cdot\|$. *Suppose* $\mu(x,\cdot) ! E \times \sigma_2 \to H$ *is a transition-measure. Under some restrictions on* $\mu(x,\cdot)$, *the set function* $m(A) = \int \mu(x, A) P(dx),\ A \in \sigma_2$ *is countably additive, where the integral refers to Bochner integral (cf.[3]). The object here is to investigate into the perfectness of the measure m.*

2. NOTATIONS.

Throughout our discussions, H will denote a real Banach-Space norm on which is denoted by $\| \cdot \|$ and Zero element of which is denoted by θ. Let (Ω_1, σ_1), (Ω_2, σ_2) be two measurable spaces, $\sigma_1 \otimes \sigma_2$ will denote the smallest σ-algebra on $\Omega_1 \times \Omega_2$ containing the set $(A : A = A_1 \times A_2$, where $A_i \in \sigma_i$, for all $i \in \{1, 2\})$. For each $A \in \sigma_1 \otimes \sigma_2$, we denote, projection of A on Ω_1 $= (x \in \Omega_1 :$ there exists $y \in \Omega_2$ such that, $(x, y) \in A)$, $A(x)$ (where $x \in \Omega_1) = (y \in \Omega_2 : (x, y) \in A)$.

The definition of projection of A on Ω_2 will be obtained by replacing in the definition of projection of A on Ω_1 1 by 2 and 2 by 1. Here ϕ will denote empty set.

Let $m : \sigma_2 \to H$ be a measure, where convergence in H is to be understood as convergence in norm. Let σ_2' be a sub-σ-algebra of σ_2. m is said to be of bounded variation on σ_2', if there exists $K \in (0, \infty)$ such that, for every sequence $(A_1', A_2', \cdots)$ of disjoint σ_2'-measurable sets, $\sum_{i=1}^{\infty} \| m(A_i') \| \leq K$. Let $P : \sigma_1 \to [0, 1]$ be a probability-measure. Let $A \subset \Omega_1$.

Let $$P_1(A) = \inf_{B \in \sigma_1 \text{ and } A \subset B} P(B).$$ *P_1 is called the outer-measure induced by P on the set of all subsets of Ω_1. P is said to be perfect, if for every σ_1-measurable function $f : \Omega_1 \to (-\infty, \infty)$, there is a Borel subset E of $f(\Omega_1)$ such that, $P(f^{-1}(E)) = 1$. A function $\mu(x, \cdot) : \Omega_1 \times \sigma_2 \to H$ is called a H-valued transition-measure on (Ω_1, σ_1, P) if for every $A \in \sigma_1 \otimes \sigma_2$, the function g_A (where $g_A(x) = \mu(x, A(x))$, $x \in \Omega_1$) is strongly measurable with respect to σ_1 and Bochner integrable with respect to P, for every countably generated sub-σ-algebra*

σ_2' of σ_2, there exists $N(\sigma_2') \in \sigma_1$ such that, $P(N(\sigma_2'))=1$ and for each $x \in N(\sigma_2')$, restriction of $\mu(x,\cdot)$ to σ_2' is a measure of bounded variation and further, the sequence $(g_A' : A \in \sigma_1 \otimes \sigma_2)$ (for all $A \in \sigma_1 \otimes \sigma_2$ and $x \in \Omega_1$, $g_A'(x) = \|\mu(x, A(x))\|$) is uniformly integrable on (Ω_1, σ_1, P). If $A \in \sigma_1$, I_A denotes characterestic function of the set A.

Throughout our subsequent discussions, (Ω_3, σ_3) will denote a measurable space, where the σ-algebra σ_3 is generated by a nonempty compact subclass D of σ_3 (cf. [5], p.7) such that, $\phi \in D$, $\Omega_3 \in D$, D is closed under the operation of taking finite union and countable intersection and for each $A \in D$, there is a sequence $\{E(i,A), i=1,2,\cdots\} \subset D$ such that,
$$A^c = \bigcup_{i=1}^{\infty} E(i,A) \text{ . Let}$$

(2.1) $T_1 = (A:$ for each integer $i \geq 1$, there is an integer $n_i \geq 1$ and a sequence $(A(i,j), B(i,j), j=1, \cdots n_i)$ such that, each $A(i,j)$ is a closed subset of [0,1], each $B(i,j) \in D$ and $A = \bigcap_{i=1}^{\infty} (\bigcup_{j=1}^{n_i} (A(i,j) \times B(i,j))))$

Given $A \in T_1$, one can easily show that, for each $x \in [0,1]$, $A(x) \in D$, (Projection of A on [0,1]) is a compact subset of [0,1] and (Projection of A on Ω_3) $\in D$.

3. THE MAIN THEOREM AND SOME OF ITS CONSEQUENCES.

To discuss our main theorem, we need the following lemma, the proof of which uses an argument due to Blackwell and Ryll-Nardzewski (cf. [1]) very closely.

LEMMA 1. Let F = Borel-σ-algebra on [0,1]. Let P be a Probability-measure on ([0,1],F) and let P_1 denote the outer-measure induced by P on the set of all subsets of [0,1]. Let $\mu(x,\cdot) : F \times \sigma_3 \rightarrow H$ be a H-valued transition-measure on ([0,1],F,P). Let $A \in F \otimes \sigma_3$, $\delta \in (0,\infty)$. Then there exists E $\in T_1$ (cf. section 2) such that, $E \subset A$ and
$$\sup_{E_1 \in F \otimes \sigma_3} \int \|\mu(x, E_1(x) \cap (A(x) - E(x)))\| P(dx) \leq \delta .$$

Proof. (1) Consider the set,

(1.1) $T = (A \in F \otimes \sigma_3 :$ given $\delta \in (0,\infty)$, there is a subset $E(A,\delta)$ of A such that, $E(A,\delta) \in T_1$ and further,

$$\sup_{E_1 \in F \otimes \sigma_3} \int \| \mu(x, E_1(x) \cap (A(x) - (E(A,\delta))(x))) \| P(dx) \leq \delta).$$

If $A_1 \in F$, $B_1 \in D$, $A_1 \times B_1 \in T$. Further, if $(E_1, \cdots E_n)$ be a finite number of disjoint elements of T, $\bigcup_{i=1}^{n} E_i \in T$. Let $\{E(1,j), j=1,2,\cdots\}$ be a nonincreasing sequence of elements of T and let $E_0 = \bigcap_{j=1}^{\infty} E(1,j)$. For each integer $n \geq 1$, choose $\alpha_{1n} \in T_1$ such that, $\alpha_{1n} \subset E(1,n)$ and

$$\sup_{E_1 \in F \otimes \sigma_3} \int \| \mu(x, E_1(x) \cap ((E(1,n))(x) - \alpha_{1n}(x))) \| P(dx) \leq 2^{-n} \delta.$$

Let $\alpha_1 = \bigcap_{n=1}^{\infty} \alpha_{1n}$. $\alpha_1 \in T$, $\alpha_1 \subset E_0$. Let $E_1 \in F \otimes \sigma_3$. We have that, $\int \| \mu(x, E_1(x) \cap (E_0(x) - \alpha_1(x))) \| P(dx)$

$$\leq \sum_{n=1}^{\infty} \int \| \mu(x, E_1(x) \cap E_0(x) \cap (\bigcap_{i=1}^{n-1} \alpha_{1i})(x) \cap ((E(1,n))(x) - \alpha_{1n}(x)) \| P(dx)$$

$\leq \delta$, where if n = 1, $(\bigcap_{i=1}^{n-1} \alpha_{1i})(x)$ should be interpreted to be Ω_3. Thus, $E_0 \in T$.

(2) Suppose $\{E(2,n), n=1,2,\cdots\}$ is a nondecreasing sequence of elements of T. Let $\delta \in (0,\infty)$. We shall prove that,

(2.1) $$\limsup_{n \to \infty} \sup_{E_1 \in F \otimes \sigma_3} \int \| \mu(x, E_1(x) \cap (E'(x) - (E(2,n))(x))) \| P(dx) \leq \delta,$$

where, $E' = \bigcup_{n=1}^{\infty} E_{2n}$.

Suppose this is not the case, choose a sequence $(n_1 < n_2 < \cdots) \subset (1,2,3,\cdots)$ and a sequence $(F_{2i}, i=1,2,\cdots) \subset F \otimes \sigma_3$ such that, for all $i \geq 1$,

(2.2) $$\int \| \mu(x, F_{2i}(x) \cap (E'(x) - (E(2,n_i))(x))) \| P(dx) \geq \delta/2$$

Let F_0 be a countably generated sub-σ-algebra of σ_3 such that, each element of the sequence $(E', E(2,n_i), F_{2i},\ i=1,2,\cdots)$ is measurable with respect to $F \otimes \sigma_0$. Choose a F-measurable N such that, P(N) = 1, for each $x \in N$, restriction of the measure $\mu(x,\cdot)$ to F_0 is of bounded variation on F_0 and further, for each $A \in F_0$, restriction of the function $\|\mu(\cdot, A)\|$ to N is $F \cap N$- measurable. Let $(F_{0n}, n=1,2,\cdots)$ be an increasing sequence of finite sub-σ-algebras of F_0 such that, $\bigcup_{n=1}^{\infty} F_{0n}$ generates F_0. For each integer $n \geq 1$, let $(A(n,1),\cdots, A(n,K(n)))$ be an enumeration of all the distinct set-theoretic atoms of F_{0n}. Let $x \in N$, $A \in F_0$. Define

$$(2.3)\quad \mu'(x,A) = \limsup_{n \to \infty} \sum_{i=1}^{K(n)} \|\mu(x, A \cap A(n,i))\|$$

For all $A \in F_0$, $\mu'(\cdot, A) : N \to [0,\infty)$ is a $F \cap N$-measurable function. Further, for all $A \in F_0$ and $x \in N$,

$$(2.4)\quad \mu'(x,A) = \operatorname{Sup} \sum_{i=1}^{n} \|\mu(x, B_i)\|$$

where, the supremum ranges over all possible finite F_0-measurable partitions $(B_1, \cdots B_n)$ of A. Choose $\delta_1 \in (0,\infty)$ such that, $A \in F$ and $P(A) \leq \delta_1$ always implies that

$$(2.5)\quad \operatorname*{Sup}_{E \in F \otimes \sigma_3} \int_A \|\mu(x, E(x))\|\, P(dx) \leq 8^{-1}\delta$$

Choose $A_0 \in F$ and $K \in (0,\infty)$ such that, $A_0 \subset N$, $P(A_0^c) \leq \delta_1$, and for all $x \in A_0$, $\mu'(x, \Omega_3) \leq K$. Choose an integer $i_0 \geq 1$ such that, for all $i \geq i_0$, $\int_{A_0} \mu'(x, E'(x) - (E(2,n_i))(x))\ .\ P(dx) \leq \delta/4$. This equation together with (2.5) contradicts (2.2). Thus (2.1) is proved. Thus T is a monotone class (cf. [8], p.13). Let

$$(2.6)\quad T_2 = (A \in \sigma_3 : \text{ for each integer } n \geq 1, \text{ there are sets } A_{1n} \in D, A_{2n} \in D \text{ such that, } A = \bigcup_{n=1}^{\infty} A_{1n},\ A^c = \bigcup_{n=1}^{\infty} A_{2n})$$

T_2 is an algebra generating σ_3. By what has already been proved, T contains an algebra generating $F \otimes \sigma_3$ and thus by a known theorem (cf. [8], p. 13), $T = F \otimes \sigma_3$. Thus the theorem is proved.

In theorem 1 that is being stated below, F will denote Borel-σ-algebra on [0,1], P will denote an arbitrary probability-measure on ([0,1],F), P_1 *will denote the outer-measure induced by P on the set of all subsets of [0,1],* (Ω_2, σ_2) *will denote an arbitrary measurable space,* f *will denote an arbitrary function which is measurable from* (Ω_2, σ_2) *to* (Ω_3, σ_3), $\mu(x, \cdot) : [0,1] \times \sigma_2 \to H$ *will denote a H-valued transition-measure on ([0,1], F, P). For all* $A \in \sigma_2$, *denote* $m(A) = \int \| \mu(x, A) \| P(dx)$. *In this context, we have the following theorem which appears to be new.*

THEOREM 1. The following two conditions are equivalent :

(1) there exists a $F \otimes \sigma_3$*- measurable set* N *such that,*

$P_1([0,1]$ *- projection of N) = 1, for all* $x \in ([0,1]$*- Projection of N),* $N(x) \subset f(\Omega_2)$ *and further,*

$$\sup_{A \in \sigma_3} P_1 \left[x \in ([0,1] \text{- projection of } N) : \| \mu(x, f^{-1}(A, \cap (N(x))^c)) \| > 0 \right] = 0,$$

(2) there exists a σ_3*-measurable set* N_1 *such that,* $N_1 \subset f(\Omega_2)$ *and for each* σ_3*-measurable subset E of* N_1^c, $m(f^{-1}(E)) = 0$

Proof. Suppose condition (1) stated in the theorem holds. Let N be the set whose existence has been ensured by condition (1). By lemma 1, there exists $N_\delta \subset N$ *such that,* $N_\delta \in T_1$,

$$\sup_{E_1 \in F \otimes \sigma_3} \int \| \mu(x, f^{-1}(E_1(x) \cap (N(x) - (N_\delta)(x)))) \| P(dx) \leq \delta.$$

Let $E_\delta = \Omega_3 -$ *projection of* N_δ, $E_\delta \subset f(\Omega_2)$ *and* $E_\delta \in \sigma_3$. *Let* $G \in \sigma_3$ *be such that,* $G \subset (E_\delta)^c$.

$$\int \| \mu(x, f^{-1}(G)) \| P(dx) = \int \| \mu(x, f^{-1}(G \cap (N(x) - (N_\delta)(x)))) \| P(dx)$$

$\leq \delta$. *Thus for each integer n ≥ 1, there exists a* σ_3*-measurable subset* B_n *of* $f(\Omega_2)$ *such that, for every* σ_3 *-measurable subset* B_n' *of* B_n^c,

$\int \| \mu(x, f^{-1}(B_n')) \| P(dx) \leq n^{-1}$. *The set* $\bigcup_{n=1}^{\infty} B_n$ *serves the purpose of the set* N_1 *given by condition 2 of theorem 1. The other part of the theorem is trivial.*

In our subsequent discussions, (E,F,P) will denote a perfect probability-space, (Ω_2, σ_2) *will denote an arbitrary measurable space,* $\mu(x,\cdot) : E \times \sigma_2 \to H$ *will denote a H-valued transition-measure on (E,F,P). In our subsequent discussions, we shall assume that, the* σ*-algebra* σ_3 *is countably generated and for each* $x \in \Omega_3$, $\{x\} \in \sigma_3$. *In our subsequent discussions,* P_1 *will denote the outer-measure induced by P on the set of all subsets of E.*

Let $\{(E_i, F_i), i = 1, 2, 3\}$ *be three measurable spaces, where* F_3 *is countably generated. Let* $\{A_1, A_2, \cdots\} \subset F_3$ *be a countable generator of* F_3. *Let* $h(x) = \sum_{n \geq 1}^{\infty} (2/3)^n I_{A_n}(x)$, *for all* $x \in E_3$. h *is called a Marczewski function corresponding to* F_3. *Let* $m : F_1 \otimes F_2 \to H$ *be a measure, where convergence in* H *is to be understood as convergence in norm. Let* f *be measurable from* (E_2, F_2) *to* (Ω_3, σ_3). m *is called* (f, F_1, F_2, σ_3)*-perfect, if there exists a* σ_3*-measurable subset B of* $f(E_2)$ *such that, for all* $A \in F_1$ *and for every* σ_3*-measurable subset* B' *of* B^c, $\| m(A \times f^{-1}(B')) \| = 0$. m *is called* (F_1, F_2, σ_3) *- perfect, if for each function* f *which is measurable from* (E_2, F_2) *to* (Ω_3, σ_3), m *is* (f, F_1, F_2, σ_3) *— perfect.*

In this context, we present the following generalisation of a theorem of Ramachandran (1973).

COROLLARY 1. Let f *be a function which is measurable from* (Ω_2, σ_2) *to* (Ω_3, σ_3) *and which is such that, there exists* $N \in F$ *such that, P(N) = 1 and for each* $x \in N$, *there is a countable subset* S_x *of* $f(\Omega_2)$ *such that, for each* σ_3*-measurable subset* E_0 *of* $(S_x)^c$, $\| \mu(x, f^{-1}(E_0)) \| = 0$.

Let $m(A) = \int \mu(x, A(x)) P(dx), A \in F \otimes \sigma_2$. *Then* m *is* $(f, F, \sigma_2, \sigma_3)$ *- perfect.*

Proof. (1) Choose a countably generated sub-σ-algebra F_2 of σ_2, a countably generated sub-σ-algebra F_1 of F and a set $N_1 \in F_1$ such that, $P(N_1)=1$, for each $x \in N_1$, restriction of $\mu(x,\cdot)$ to F_2 is a measure of bounded variation, f is measurable from (Ω_2, F_2) to (Ω_3, σ_3) and further, for every $A \in F_2$, the function $g_A(x) = \mu(x, A)$, $x \in N_1$ is measurable from $(N_1, F_1 \cap N_1)$ to (H, B) (where B = Borel-σ-algebra on H), g_A is strongly measurable with respect to $F_1 \cap N_1$ as well as Bochner integrable with respect to the restriction of P to $F_1 \cap N_1$. Let $(A_1, A_2, \cdots) \subset F_1$ be a countable generator of F_1. Let $h_1(x) = \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^n I_{A_n}(x)$, for all $x \in E$. Choose a Borel subset N_2 of $h_1(N_1)$ such that, $P(h_1^{-1}(N_2)) = 1$. Let $F' =$ Borel-σ-algebra on $[0,1]$. For all $A \in F'$, define $P_1(A) = P(h_1^{-1}(A))$. Let $x \in N_2$. Let $A(x)$ be the unique set-theoretic atom of F_1 such that, for all $y \in A(x)$, $h_1(y) = x$. Let $y \in A(x)$, $A_1 \in \sigma_3$. Define $\mu_2(x, A_1) = \mu(y, f^{-1}(A_1))$. If $x \in N_2^c$, define $\mu_2(x, A_1) = \theta$.

(2) Let $\{\sigma_{3n}, n = 1, 2, \cdots\}$ be a nondecreasing sequence of finite sub-σ-algebras of σ_3 such that, for each integer $n \geq 1$ and for each set-theoretic atom A of σ_{3n}, A is expressible as union of the elements of a set $\{D(n, A), n = 1, 2, \cdots\} \subset D$ (cf. equation (2.6) in the proof of lemma 1) and further, $\sigma_0 = \bigcup_{n=1}^{\infty} \sigma_{3n}$ is an algebra generating σ_3. For each integer $n \geq 1$, let $(A(n,1), \cdots\cdots A(n, K_n))$ be an enumeration of all the distinct set-theoretic atoms of σ_{3n}. Let $D_3 = (A' \in \sigma_3 :$ either $A' = \phi$ or A' is expressible as union of a finite number of elements of the set $\{D(t, A(n,i)), t = 1, 2, \cdots, i = 1, \cdots K(n), n = 1, 2, \cdots\})$. For all $x \in [0,1]$ and $A_1 \in \sigma_3$, define $\mu_3(x, A_1) = \operatorname{Sup} \sum_{i=1}^{t} \| \mu_2(x, \alpha_i) \|$, where the supremum ranges over all possible finite partitions $(\alpha_1, \cdots \alpha_t)$ of A_1. For each $x \in [0,1]$, $\mu_3(x, \cdot) : \sigma_3 \to [0, \infty)$ is a measure. Let $\delta \in (0, \infty)$, $\delta_1 \in (0, \infty)$.

For each $x \in N_2$ and $A \in \sigma_0$, there is a set $D'(x, A) \subset A$ such that, $D'(x, A) \in D_\delta$ and $\mu_3(x, A - D'(x, A)) \le \delta_1$ and consequently,

$$\sup_{E_0 \in \sigma_0} \| \mu_2(x, E_0 \cap (A - D'(x, A))) \| \le \delta_1$$

D_δ is a countable subset of D. Choose $\delta_2 \in (0, \infty)$ such that $A \in F_1$ and $P_1(A) \le \delta_2$ always implies that, for all $E_0 \in F' \otimes \sigma_3$,

(2.1)
$$\int_A \| \mu_2(x, E_0(x)) \| \, P_1(dx) \le \delta$$

Using some of the ideas presented above, one can show that, for each integer n ⩾ 1, there exists a F'-measurable subset $A(n, \delta_2)$ of N_2 such that, $P_1(A(n, \delta_2)) \geqslant 1 - 2^{-n} \delta_2$ and $\forall i = 1, \cdots K(n)$, there exists a subset $E(n, i)$ of $A(n, i)$ such that $E(n, i) \in D$ and

(2.2)
$$\sup_{\substack{E_0 \in \sigma_0 \text{ and } x \in A(n, \delta_2)}} \left\| \mu_2(x, E_0 \cap (A(n, i) - E(n, i))) \right\| \le 2^{-(n+i+K(n))}.$$

Let $\alpha(\delta) = \bigcap_{n=1}^{\infty} \left(\bigcup_{i=1}^{K(n)} E(n, i) \right)$, $A(\delta_2) = \bigcap_{n=1}^{\infty} A(n, \delta_2)$. One can easily show that, for all $x \in A(\delta_2)$, $\sup_{\alpha_0 \in \sigma_3} \| \mu_2(x, \alpha_0 \cap (\alpha(\delta))^c) \| \le \delta$. Since for all $x \in \Omega_3$, $\{x\} \in \sigma_3$,

(2.3)
$$[(x, y) : x \in \alpha(\delta) \text{ and } x = y] = \bigcap_{n=1}^{\infty} \left[\bigcup_{i=1}^{K(n)} (\alpha(\delta) \cap A(n, i) \times \alpha(\delta) \cap A(n, i)) \right] \in \sigma_3 \otimes \sigma_3.$$

Using this fact and imitating some known steps (cf. [6] and [3], p.52), we get that,

(2.4)
$$[(x, y) \in A(\delta_2) \times \alpha(\delta) : \| \mu_2(x, \{y\}) \| > 0] \in F' \otimes \sigma_3.$$

Using this and theorem 1, one can easily show that, there exists a σ_3-measurable subset $B(\delta)$ of $f(f^{-1}(\alpha(\delta)))$ such that, for every σ_3-measurable subset α_0 of $(B(\delta))^c \cap \alpha(\delta)$,

$$\int_{h_1^{-1}(A(\delta_2))} \| \mu(x, f^{-1}(\alpha_0)) \| P(dx) = 0 \tag{2.5}$$

Using equations (2.1), (2.5) and the fact that, $\sup\limits_{x \in A(\delta_2) \text{ and } \alpha_0 \in \sigma_3} \| \mu_2(x, \alpha_0 \cap (\alpha(\delta))^c) \|$ $\leq \delta$, *we get that, for every* σ_3*-measurable subset* α_0 *of* $(B(\delta))^c$,

$$\int \| \mu(x, f^{-1}(\alpha_0)) \| P(dx) \leq \delta + \int_{h_1^{-1}(A\delta_2)} \| \mu(x, f^{-1}(\alpha_0 \cap (\alpha(\delta))^c)) \| P(dx) \leq 2\delta. \tag{2.6}$$

Thus for every integer $n \geq 1$, *there exists a* σ_3*-measurable subset* B_n of $f(\Omega_2)$ *such that, for every* σ_3*-measurable subset* α_0 *of* B_n^c, $\int \| \mu(x, f^{-1}(\alpha_0)) \| P(dx) = 0$. *Let* $B_0 = \bigcup\limits_{n=1}^{\infty} B_n$. *The set* B_0 *serves our purpose.*

The following result. appears to be new.

THEOREM 2. Let σ_0 *denote Borel-*σ*-algebra on* $(-\infty, \infty)$. *Let*

$$m(A) = \int \mu(x, A(x)) P(dx), \quad A \in F \otimes \sigma_2$$

m *is* (F, σ_2, σ_0)*- perfect if and only if, for every countably generated sub-*σ*-algebra* σ_2' *of* σ_2, *there exists a Marczewski-function* $h_{\sigma_2'}$ *corresponding to* σ_2' *and a* $F \otimes \sigma_0$*- measurable set* $N_{\sigma_2'}$ *such that,* P_1 *(E-projection of* $N_{\sigma_2'}$ *) =1, for all* $x \in E$, $N_{\sigma_2'}(x) \subset h_{\sigma_2'}(\Omega_2)$ *and*

$$\sup_{A \in \sigma_0} P_1 \left[x \in E - \text{Projection of } N_{\sigma_2'} : \| \mu(x, h_{\sigma_2'}^{-1}(A) \cap (N_{\sigma_2'}(x))^c) \| > 0 \right] = 0$$

Proof. "$\underline{\underline{\textit{If}}}$*". Let B=Borel-*$\sigma$*-algebra on H. Let* f *be a function which is measurable from* (Ω_2, σ_2) *to* $((-\infty, \infty), \sigma_0)$. *Let* F_2 *be a countably generated sub-*σ*-algebra of* σ_2 *such that,* f *is measurable from* (Ω_2, F_2) *to* $((-\infty, \infty), \sigma_0)$. *Let* h_{F_2} *be a Marczewski-function corresponding to* F_2 *and let* N_{F_2} *be a* $F \otimes \sigma_0$ *- measurable function whose existence has been ensured by the condition of theorem 2. Let* F_1 *be a countably generated sub-*σ*-algebra of F and let* $N_1 \in F_1$ *be such that,*

$P(N_1)=1$, $N_1 \subset$ (E-projection of N_{F_2})

(cf. lemma 1), for every $x \in N_1$, $\mu(x,\cdot)$ is of bounded variation on (Ω_2, F_2), for every $A \in F_2$, the function $g_A(x) = \mu(x,A)$, $x \in N_1$ is measurable from $(N_1, F_1 \cap N_1)$ to (H, B) and further, the function g_A is strongly measurable with respect to $F_1 \cap N_1$ as well as Bochner-integrable with respect to the restriction of P to $F_1 \cap N_1$. To prove this, one has to use the known fact that (cf.[7], p.4), F_1 and N_1 can be so chosen that, there exists a Marczewski-function h corresponding to F_1 such that, $h(N_1) \in \sigma_0$ and in this case, for all $A \in F_1 \cap N_1$, $h(A)$ is a Borel subset of [0,1]. For each integer $n \geq 1$, choose a $F_1 \cap N_1$-measurable nonempty set A_n and $\delta_n \in (0,\infty)$ such that, for all $x \in A_n$,

$\sup \sum_{i=1}^{t} \|\mu(x, B_i)\| \leq \delta_n$, where the supremum ranges over all possible F_2-measurable finite partitions $(B_1, \cdots B_t)$ of Ω_2 and further for all $A \in F_1 \otimes F_2$, $\int_{A_n^c} \|\mu(x, A(x))\| P(dx) \leq n^{-1}$. Let $n \in (1,2,\cdots)$.

Let $N_{2n} = [(h(x),y) : (x,y) \in (A_n \times (-\infty,\infty)) \cap N_{F_2}]$.

Let F_3 = Borel-σ-algebra on [0,1]. Using a known theorem (cf.[7],p.4), one can easily show that, $N_{2n} \in F_3 \otimes \sigma_0$. Let $x \in h(A_n)$, $E_0 \in F_2$.

Choose a unique set-theoretic atom $B(x)$ of F_1 which is such that, $B(x) \subset A_n$ and for all $y \in B(x)$, $h(y) = x$. Let $y \in B(x)$. Define $\mu_1(x, E_0) = \mu(y, E_0)$. Using theorem 1, one can easily show that, there exists a nonempty Borel subset B_n of $h_{F_2}(\Omega_2)$ such that, for every σ_0-measurable subset B_n' of B_n^c

(1) $$\int_{h(A_n)} \|\mu_1(x, B_n')\| P_2(dx) = 0$$

where for all $A \in F_3$, $P_2(A) = P(h^{-1}(A))$. Using a known theorem (cf. [7], p.4), one can easily show that, for every F_2-measurable subset B_n' of $h_{F_2}^{-1}(B_n^c)$,

(2) $$\int_{A_n} \|\mu(x, B_n')\| \, P(dx) = 0$$

Let $x \in B_n'$. There is a unique set-theoretic atom $\alpha(x)$ of such that, for all $y \in \alpha(x)$, $h_{F_2}(y) = x$. Let $y \in \alpha(x)$. Define $g(x) = f(y)$. g is measurable from $(B_n, F_3 \cap B_n)$ to $((-\infty, \infty), \sigma_0)$.

For all $x \in h(A_n)$ and $\alpha_0 \in \sigma_0 \cap B_n$, define $\mu_2(x, \alpha_0)$ $= \operatorname{Sup} \sum_{i=1}^{t} \|\mu_1(x, \alpha_i)\|$, where the supremum ranges over all possible σ_0-measurable finite partitions $(\alpha_1, \ldots \alpha_t)$ of α_0. Using the fact that, any bounded positive measure defined on Borel subsets of a Borel subset of [0,1] is perfect, one can easily show that, there exists a Borel subset β_n of $g(B_n)$ such that, $\int_{h(A_n)} \mu_2(x, B_n - g^{-1}(\beta_n)) \, P_2(dx) = 0$

Using this and a known fact (cf.[7], p.4), one can easily show that, for every Borel subset β_n' of β_n^c,

(3) $$\int_{A_n} \|\mu(x, f^{-1}(\beta_n'))\| \, P(dx) = 0$$

By (3), for every Borel subset β_n' of β_n^c,

$$\int \|\mu(x, f^{-1}(\beta_n'))\| \, P(dx) \leq n^{-1}.$$

The set $\beta_0 = \bigcup_{n=1}^{\infty} \beta_n$ has the property that, $\beta_0 \in \sigma_0$, $\beta_0 \subset f(\Omega_2)$ and further, for every Borel subset β_0' of β_0^c,

$$\int \|\mu(x, f^{-1}(\beta_0'))\| \, P(dx) = 0$$

"Only if". This part is a trivial consequence of a known result (cf.[2]).

REFERENCES

1. BLACKWELL, D and RYLL-NARDZEWSKI, C. (1963) : "Non-existence of everywhere proper conditional distributions", Ann. Math. Statist., 34, 223-225.

2. GHOSH, S. (1983) : "On a theorem in Real analysis", Bulletin of the Calcutta Mathematical Society, 75, 369-371.

3. GHOSH, S. (1992) : "On some theorem on perfect probabilities and the strong law of large numbers", Proceedings of the international A.M.S.E. Conference held in New Delhi, India during December 9-11, 1991, Vol.I, P. 47-59.

4. GHOSH, S. (1993) : " A version of Borel-equivalence theorem and some of its consequences", Proceedings of the international A.M.S.E. Conference held in Calcutta, India during December 7-9, 1992, Vol I, p. 39-49.

5. GHOSH, S. (1993) : "On preservation of compact approximation property by thick subsets of a probability-space and some other remarks", Proceedings of the international A.M.S.E. Conference held in Calcutta, India during December 7-9, 1992, Vol. I, p. 15-26.

6. RAMACHANDRAN, D. (1974) : "Mixture of perfect Probability-measures", Ann. Probability 2, 495.

7. RAMACHANDRAN, D. (1979) : "Perfect Measures", Part I, I.S.I. Lecture Notes, No. 5, Indian Statistical Institute, Calcutta.

8. NEVEU, J. (1965) : "Mathematical Foundations of the calculus of Probability", Holden Day, San-Fransisco.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 39-52

ON GAMMA-FUNCTIONS AND SUM OF INDEPENDENT RANDOM VARIABLES

By

S. GHOSH

Department of Statistics, Presidency College, Calcutta-700 073

INDIA

1. *ABSTRACT AND INTRODUCTION*

If α, β *be two real numbers such that* $\alpha > 0$, *several properties of the function* $\Gamma(\alpha + i\beta)$ *can be studied. In this paper, we have first of all obtained an approximate expression of* $\Gamma(\alpha + i\beta)$, *when* α *is large. We have then obtained a result which has bearing on summability of independent random variables.*

2. *AN APPROXIMATE EXPRESSION OF GAMMA FUNCTIONS OF A COMPLEX VARIABLE.*

THEOREM 1. Let α, β *be real numbers such that,*

$\alpha \in (5, \infty)$. *Then* $|\Gamma(\alpha + i\beta)| > 0$ *and,*

$$\Big| \log_e |\Gamma(\alpha + i\beta)| - \Big[\log_e \sqrt{2\pi} + \tfrac{1}{2}(\alpha - \tfrac{1}{2}) \log_e (\alpha^2 + \beta^2) - \beta^2 \sum_{t \in (\alpha, \alpha+1, \alpha+2, \ldots)} (t^2 + \beta^2)^{-1} + \frac{\beta^2}{2} (\beta^2 + \alpha^2)^{-1} - \alpha \Big] \Big| \leq 6(\alpha - 4)^{-1} + 8(\alpha - 1)^{-2} + 2\Big(\frac{2}{\alpha} + \log_e \frac{\alpha + 2}{\alpha - 1}\Big)$$

Proof. (1) Let n *be an integer which is* $> \sqrt{3}|\beta| + 5$ *Let* $z = \alpha + i\beta$

Let

$$(1.1) \quad S_1(\alpha, \beta, n) = \log_e |z(z+1) \cdots (z+n)|.$$

Then,

$$S_1(\alpha, \beta, n) = 2^{-1} \sum_{t=1}^{n} (t - \alpha - n + \alpha + n) \log_e \left[\frac{(\alpha + n + 1 - t)^2 + \beta^2}{(\alpha + n - t)^2 + \beta^2} \right] + 2^{-1}(n+1) \log_e (\alpha^2 + \beta^2)$$

$$= 2^{-1}(\alpha + n) \left[\log_e \{ (\alpha + n)^2 + \beta^2 \} - \log_e (\alpha^2 + \beta^2) \right]$$

$$+ 2^{-1}\sum_{t=1}^{n}(t-\alpha-n)\log_e\left[1+\frac{2(\alpha+n-t)+1}{(\alpha+n-t)^2+\beta^2}\right] + 2^{-1}(n+1)\log_e(\alpha^2+\beta^2)$$

Since $\alpha \in (3, \infty)$, *for all* $t = 1, \cdots n$, $1+2(\alpha+n-t) < (\alpha+n-t)^2+\beta^2$

and so,

(1.2) $$S_1(\alpha,\beta,n) = S_2(\alpha,\beta,n) + S_3(\alpha,\beta,n),$$ *where,*

(1.3) $$S_2(\alpha,\beta,n) = -2^{-1}(\alpha-1)\log_e(\alpha^2+\beta^2) + 2^{-1}(\alpha+n)\log_e\{(\alpha+n)^2+\beta^2\}$$

(1.4) $$S_3(\alpha,\beta,n) = 2^{-1}\sum_{t=1}^{n}(t-\alpha-n)\log_e\left[1+\frac{1+2(\alpha+n-t)}{(\alpha+n-t)^2+\beta^2}\right]$$

$$= 2^{-1}\sum_{t=1}^{n}(t-\alpha-n)\left[\frac{1+2(\alpha+n-t)}{(\alpha+n-t)^2+\beta^2} - \frac{1}{2}\left(\frac{1+2(\alpha+n-t)}{(\alpha+n-t)^2+\beta^2}\right)^2\right]$$

$$+ \quad S_4(\alpha,\beta,n),$$ *where,*

(1.5) $$S_4(\alpha,\beta,n) = 2^{-1}\sum_{t=1}^{n}(t-\alpha-n)\left[\sum_{K=3}^{\infty}\frac{(-1)^{K-1}}{K}\left(\frac{1+2(\alpha+n-t)}{(\alpha+n-t)^2+\beta^2}\right)^K\right]$$

Using C.R.inequality (cf. [6], p.155), one can show that,

(1.6) $$\left|S_4(\alpha,\beta,n)\right| \le 6^{-1}\sum_{t=1}^{n}\frac{(\alpha+n-t)(1+2(\alpha+n-t))^3}{[(\alpha+n-t)^2+\beta^2]^3}\left(1-\frac{1+2(\alpha+n-t)}{(\alpha+n-t)^2+\beta^2}\right)^{-1}$$

$$\le 6^{-1}\sum_{t=1}^{n}\frac{4(\alpha+n-t)(1+8(\alpha+n-t)^2)}{((\alpha+n-t)^2+\beta^2)^2} \times \left[(\alpha+n-t)^2+\beta^2-1-2(\alpha+n-t)\right]^{-1}$$

$$\le 6^{-1}\sum_{t\in(\alpha,\alpha+1,\alpha+2,\cdots)}(t-3)^{-2} \le 6^{-1}\int_{\alpha-4}^{\infty}x^{-2}\,dx = 6\,(\alpha-4)^{-1}$$

Using (1.4), one can show that,

(1.7) $$S_3(\alpha,\beta,n) = S_4(\alpha,\beta,n) + S_5(\alpha,\beta,n) + S_6(\alpha,\beta,n),$$ *where,* (1.8) $$S_5(\alpha,\beta,n) = -n + \beta^2 \sum_{t=1}^{n} \frac{1}{(\alpha+n-t)^2+\beta^2} - \frac{1}{2}\sum_{t=1}^{n} \frac{\alpha+n-t}{(\alpha+n-t)^2+\beta^2} + \sum_{t=1}^{n} \frac{(\alpha+n-t)^3}{[(\alpha+n-t)^2+\beta^2]^2}$$

(1.9) $$S_6(\alpha,\beta,n) = \frac{1}{4}\sum_{t=1}^{n} (\alpha+n-t) \times \frac{1+4(\alpha+n-t)}{\{(\alpha+n-t)^2+\beta^2\}^2}$$

We have that,

(1.10) $$|S_6(\alpha,\beta,n)| \le \int_{\alpha-1}^{\infty} (4^{-1}x^{-3} + x^{-2})\,dx = 8^{-1}(\alpha-1)^{-2} + (\alpha-1)^{-1}$$

(2) Let $x \in (0,\infty)$, $j \in \{1,2\}$. *Define* $f_j(x) = x^{2j-1}(x^2+\beta^2)^{-1}$ *Observe by differentiation that,* f_j *is a nondecreasing function in* $(0, \alpha_j|\beta|)$ *and is a nonincreasing function in* $(\alpha_j|\beta|, \infty)$, *where if* $j=1$, $\alpha_j = 1$ *and if* $j=2$, $\alpha_j = \sqrt{3}$. *Let*

(2.1) $$S_7(\alpha,\beta,n) = \sum_{t\in(\alpha,\alpha+1,\dots\alpha+n-1)} f_1(t)$$

Suppose $\alpha < |\beta|$. *Let* K *be the largest nonnegative integer such that,* $\alpha + K < |\beta|$. We *have that,*

(2.2) $$\int_{\alpha-1}^{\alpha+K} x(x^2+\beta^2)^{-1}dx + \int_{\alpha+K+2}^{\alpha+n-1} x(x^2+\beta^2)^{-1}dx + \sum_{t\in\{K+1,n-1\}} f_1(\alpha+t)$$
$$\le \sum_{t\in(\alpha,\alpha+1,\dots\alpha+n-1)} f_1(t)$$
$$\le \int_{\alpha}^{\alpha+K} x(x^2+\beta^2)^{-1}dx + \sum_{t\in\{K,K+1\}} f_1(\alpha+t) + \int_{\alpha+K+1}^{\alpha+n-1} x(x^2+\beta^2)^{-1}dx$$

Also,

(2.3) $$\int_{\alpha}^{\alpha+n-1} x(x^2+\beta^2)^{-1}dx = \frac{1}{2}\log_e\left[\frac{(\alpha+n-1)^2+\beta^2}{\alpha^2+\beta^2}\right]$$

Using (2.2) and (2.3) and the fact that, for all $x \in (0,\infty)$, $x(x^2+\beta^2)^{-1} \le x^{-1}$, *one can easily show that, if* $\alpha < |\beta|$,

(2.4) $$\left| S_7(\alpha,\beta,n) - \frac{1}{2}\log_e\left[\frac{(\alpha+n-1)^2+\beta^2}{\alpha^2+\beta^2}\right] \right|$$

$$\leq \frac{2}{\alpha} + \int_{\alpha-1}^{\alpha} \frac{x}{x^2+\beta^2}\,dx + \int_{\alpha+k}^{\alpha+k+2} \frac{x}{x^2+\beta^2}\,dx \leq \frac{2}{\alpha} + \log_e \frac{\alpha+2}{\alpha-1}$$

Suppose now, $|\beta| \leq \alpha$. *Then,*

(2.5)
$$\sum_{t\in(\alpha,\alpha+1,\dots\alpha+n-1)} f_1(t)$$

$$\leq \frac{\alpha}{\alpha^2+\beta^2} + \int_{\alpha}^{\alpha+n-1} \frac{x}{x^2+\beta^2}\,dx \leq \frac{1}{\alpha} + \frac{1}{2}\log_e \frac{(\alpha+n-1)^2+\beta^2}{\alpha^2+\beta^2}$$

Also in the case $|\beta| \leq \alpha$,

(2.6)
$$2^{-1}\log_e \frac{(\alpha+n-1)^2+\beta^2}{\alpha^2+\beta^2} = \int_{\alpha}^{\alpha+n-1} \frac{x}{x^2+\beta^2}\,dx \leq \int_{\alpha}^{\alpha+n} \frac{x}{x^2+\beta^2}\,dx$$

$$\leq \sum_{t\in(\alpha,\alpha+1,\dots\alpha+n-1)} f_1(t)$$

Using (2.4), (2.5) and (2.6), one can show that,

(2.7)
$$\left| \sum_{t\in(\alpha,\alpha+1,\dots\alpha+n-1)} f_1(t) - \frac{1}{2}\log_e \frac{(\alpha+n-1)^2+\beta^2}{\alpha^2+\beta^2} \right| \leq \frac{2}{\alpha} + \log_e \frac{\alpha+2}{\alpha-1}$$

Proceeding in a very similar fashion, one can show that, if $\alpha < \sqrt{3}|\beta|$,

$$\left| \sum_{t\in(\alpha,\alpha+1,\dots\alpha+n-1)} t^3(t^2+\beta^2)^{-1} - \int_{\alpha}^{\alpha+n} x^3(x^2+\beta^2)^{-2}\,dx \right|$$

$$\leq \frac{2}{\alpha} + \log_e \frac{\alpha+2}{\alpha-1}$$, *whereas if* $\sqrt{3}|\beta| \leq \alpha$,

$$\int_{\alpha}^{\alpha+n-1} x^3(x^2+\beta^2)^{-2}\,dx \leq \int_{\alpha}^{\alpha+n} x^3(x^2+\beta^2)^{-2}\,dx \leq \sum_{t\in(\alpha,\alpha+1,\dots\alpha+n-1)} f_2(t)$$

$$\leq \alpha^3(\alpha^2+\beta^2)^{-2} + \int_{\alpha}^{\alpha+n-1} x^3(x^2+\beta^2)^{-2}\,dx$$

Using the above equations and the fact that, $\int_{\alpha}^{\alpha+n-1} x^3(x^2+\beta^2)^{-2}\,dx$

$$= 2^{-1}\log_e\left[\frac{(\alpha+n-1)^2+\beta^2}{\alpha^2+\beta^2}\right] - \frac{\beta^2}{2}\left[\frac{1}{\beta^2+\alpha^2} - \frac{1}{\beta^2+(\alpha+n-1)^2}\right]$$

one can easily show that,

$$(2.8)\quad \left| \sum_{t\in(\alpha,\alpha+1,\cdots\alpha+n-1)} f_2(t) - \left[2^{-1}\log_e\left\{\frac{(\alpha+n-1)^2+\beta^2}{\alpha^2+\beta^2}\right\} - \frac{\beta^2}{2}\left(\frac{1}{\beta^2+\alpha^2} - \frac{1}{\beta^2+(\alpha+n-1)^2}\right)\right]\right|$$

$$\leq \frac{2}{\alpha} + \log_e \frac{\alpha+2}{\alpha-1}$$

We have that (cf. **[2]**, p.178 and [11]),

$$(2.9)\quad \Gamma(\alpha+i\beta) = \underset{\substack{n\to\infty \\ n\in(1,2,3,\cdots)}}{\text{limit}} \frac{n!\ n^{z}}{z(z+1)\cdots\cdots(z+n)},$$

and,

$$(2.10)\quad \underset{\substack{n\to\infty \\ n\in(1,2,3,\cdots)}}{\text{limit}} \left[\log_e n! - \left\{\log_e\sqrt{2\pi} - n + (n+\tfrac{1}{2})\log_e n\right\}\right] = 0$$

Using equations (1.2), (1.3), (1.4), (1.5), (1.6), (1.7), (1.8), (1.9), (1.10), (2.1), (2.7), (2.8), (2.9) and (2.10), one can easily get at the theorem.

COROLLARY 1.

$$\underset{\substack{\text{Maximum}\left\{\frac{|\beta|}{\alpha},\ \frac{\beta^2}{\alpha},\ \alpha^{-1}\right\}\to 0 \\ \text{and}\quad \alpha>0}}{\text{limit}} \left[\log_e\left|\Gamma(\alpha+i\beta)\right| - \left[\log_e\sqrt{2\pi} + (\alpha-\tfrac{1}{2})\log_e\alpha - \alpha\right]\right] = 0$$

Proof. If $\delta\in(0,1)$ be sufficiently close to 0, then $\delta - \frac{\delta^2}{2} \leq \log_e(1+\delta) \leq \delta$. Using this, one can show that, if $\alpha>0$ be sufficiently large and $\frac{|\beta|}{\alpha} + \frac{\beta^2}{\alpha}$ be sufficiently close to 0, $(\alpha-\frac{1}{2})\log_e\left(1+\frac{\beta^2}{\alpha^2}\right)$ is sufficiently close to 0. Let $\alpha\in(2,\infty)$. $\beta^2 \sum_{t\in(\alpha,\alpha+1,\alpha+2,\cdots)} (t^2+\beta^2)^{-1} \leq \beta^2\int_{\alpha-1}^{\infty} x^{-2}\,dx = \frac{\beta^2}{\alpha-1} = \frac{\beta^2}{\alpha}\left(1+\frac{1}{\alpha-1}\right) \leq \frac{2\beta^2}{\alpha} \to 0$ as $\frac{\beta^2}{\alpha}$ gets sufficiently close to 0. Now using theorem 1, Corollary 1 can be proved very easily.

3. REMARKS ON COMPLEX SEQUENCES AND SUM OF INDEPENDENT FUNCTIONS.

THEOREM 2. Let $\mathbb{C}$ denote the set of all complex numbers, Let C_1 be a connected subset of $\mathbb{C}$ containing at least two distinct points. Let $f : C_1 \to \mathbb{C}$ be a continuous function i.e. for all $x \in C_1$, $\lim_{x_1 \in C_1 \text{ and } x_1 \to x} f(x_1) = f(x)$

We assume that, there exists a countable subset N of C_1 such that, for all

$$x \in C_1 - N, \quad \lim_{x_1 \to x,\, x_1 \neq x \text{ and } x_1 \in C_1} (x_1 - x)^{-1} (f(x_1) - f(x)) = 0 .$$

Then there is a complex number α such that, for all $x \in C_1$, $f(x) = \alpha$

Proof. Let $a \in C_1$. Let b be an element of C_1 such that $a \neq b$. Then a and b can be joined by a finite sequence of straight lines $\subset C_1$, joined end to end, each of which is either parallel to the X-axis or the Y-axis (cf.[2],p.15). Let $(\alpha_1 + i\beta_1)$, $(\alpha_2 + i\beta_1)$ be two distinct elements of C_1 such that, for all $t \in [\alpha_3, \alpha_4]$ (α_3 = Minimum $\{\alpha_1, \alpha_2\}$, α_4 = Maximum $\{\alpha_1, \alpha_2\}$), $(t + i\beta_1) \in C_1$. Suppose for all $(x + iy) \in C_1$, $f(x + iy) = u(x, y) + i\, v(x, y)$, where, u and v are real-valued functions. There is a countable subset S_{α_3, α_4} of $[\alpha_3, \alpha_4)$ such that, for all $x \in [\alpha_3, \alpha_4] - S_{\alpha_3, \alpha_4}$,

$$\lim_{h \to 0+} [u(x+h, y) - u(x, y)] = 0 = \lim_{h \to 0+} [v(x+h, y) - v(x, y)]$$

Using this and a known theorem (cf. [4], p. 42), one can easily show that, for all $x \in [\alpha_3, \alpha_4]$, $f(x + i\beta_1) = f(\alpha_1 + i\beta_1)$. Similarly if $(\alpha + i\beta_2)$, $(\alpha + i\beta_3)$ be two distinct elements of C_1 such that, for all $t \in [\beta_4, \beta_5]$ (β_4 = Minimum $\{\beta_2, \beta_3\}$, β_5 = Maximum $\{\beta_2, \beta_3\}$), $(\alpha + it) \in C_1$, then for all $t \in [\beta_4, \beta_5]$, $f(\alpha + it) = f(\alpha + i\beta_2)$

Using these facts a finite number of times, we get at theorem 2.

The following known theorem (cf.[2], p. 37) follows from theorem 2.

Let C_1 be a nonempty open connected subset of C, the set of complex numbers and let $f : C_1 \to \mathbb{C}$ be a continuous function such that, for all $z \in C_1$, the first derivative of f at z is 0. Then f is constant on C_1.

THEOREM 3. *Let* $p \in (1,\infty)$, $q \in (1,\infty)$ *be such that,* $p^{-1}+q^{-1}=1$. *Let* $(a_1, a_2, \dots)$ *be a sequence of real numbers.* $\sum_{n=1}^{\infty} |a_n|^p < \infty$, *if and only if for everysequence* $(b_1, b_2, \dots)$ *of real numbers such that,* $\sum_{n=1}^{\infty} |b_n|^q < \infty$, $\limsup_{K\to\infty} \left| \sum_{n=1}^{K} a_n b_n \right| < \infty$

Proof. *"Only if".* *Suppose* $\sum_{n=1}^{\infty} |a_n|^p < \infty$. *Let* $\{b_1, b_2, \dots\}$ *be a sequence of real numbers such that,* $\sum_{n=1}^{\infty} |b_n|^q < \infty$.

By a known result (cf. [7], p. 56), for every integer $K \geq 1$

$$\left| \sum_{n=1}^{K} a_n b_n \right| \leq \left[\sum_{n=1}^{K} |a_n|^p \right]^{p^{-1}} \left[\sum_{n=1}^{K} |b_n|^q \right]^{q^{-1}} \leq \left[\sum_{n=1}^{\infty} |a_n|^p \right]^{p^{-1}} \left[\sum_{n=1}^{\infty} |b_n|^q \right]^{q^{-1}}$$

"If." *Let* $\Omega = (1, 2, \dots)$, σ *= Set of all subsets of* Ω

Let μ *be a probability-measure on* (Ω, σ) *such that, for every integer n ⩾ 1*

$\mu(\{n\}) = 2^{-n}$. *Let* $t \in (1, \infty)$. *Define*

$L_t(\Omega, \sigma, \mu) = \{f : f : \Omega \to (-\infty, \infty)$ *is a* σ*-measurable function such that,* $\int |f(x)|^t \mu(dx) < \infty\}$.

For all $f \in L_t(\Omega, \sigma, \mu)$, *define* $\|f\|_t = \left[\int |f(x)|^t \mu(dx) \right]^{t^{-1}}$

Let $f \in L_q(\Omega, \sigma, \mu)$. *Then* $\sum_{n=1}^{\infty} \left(|f(n)| 2^{-\frac{n}{q}} \right)^q < \infty$

Then $\limsup_{K\to\infty} \left| \sum_{n=1}^{K} a_n f(n) 2^{\frac{n}{p}} 2^{-n} \right| < \infty$

For every integer K ⩾ 1, define $\Lambda_K(f) = \sum_{n=1}^{K} a_n 2^{\frac{n}{p}} f(n) 2^{-n}$

For each $f \in L_q(\Omega, \sigma, \mu)$, $\limsup_{K\to\infty} |\Lambda_K(f)| < \infty$

By the uniform boundedness theorem (cf.[8], p. 239), there exists $T \in (0, \infty)$ *such that,*

$$\sup_{K \geq 1} \left[\sup_{\substack{g \in L_q(\Omega, \sigma, \mu) \\ \text{and } \|g\|_q \leq 1}} \left| \sum_{n=1}^{K} a_n 2^{\frac{n}{p}} g(n) 2^{-n} \right| \right] \leq T$$

This shows that, for all $K \in (1, 2, \dots)$,

$$\left[\sum_{n=1}^{K} |a_n|^p 2^n 2^{-n}\right]^{p^{-1}} = \left[\sum_{n=1}^{K} |a_n|^p\right]^{p^{-1}} \le T$$

Thus, $\sum_{n=1}^{\infty} |a_n|^p < \infty$

COROLLARY 2. Let $p \in (1,\infty)$, $q \in (1,\infty)$ *be such that,* $p^{-1} + q^{-1} = 1.$ *Let* $\{a_1, a_2, \cdots\}$ *be a sequence of nonnegative real numbers such that, for every nonnegative sequence* $\{b_1, b_2, \cdots\}$ *of real numbers such that,* $\sum_{n=1}^{\infty} (b_n)^q < \infty$,

$$\sum_{n=1}^{\infty} a_n b_n < \infty \quad .$$

Then, $\sum_{n=1}^{\infty} a_n^p < \infty$

Proof Let $x \in (-\infty, \infty)$. *Define* $x^+ = \operatorname{Sup}\{x, 0\}$, $x^- = (-x)^+$

Let $\{b_1, b_2, \cdots\}$ *be a sequence of real numbers such that,*

$\sum_{n=1}^{\infty} |b_n|^q < \infty$. *For each* $n \in (1, 2, \cdots)$,

$|b_n|^q = (b_n^+)^q + (b_n^-)^q$ and so, $\sum_{n=1}^{\infty} a_n b_n^+ < \infty$, $\sum_{n=1}^{\infty} a_n b_n^- < \infty$

So, $\sum_{n=1}^{\infty} |a_n b_n| = \sum_{n=1}^{\infty} a_n (b_n^+ + b_n^-) < \infty$.

So, $\limsup_{K\to\infty} \left| \sum_{n=1}^{K} a_n b_n \right| < \infty$. *So by theorem 3, the corollary holds.*

It is known that (cf. [8], p. 98), if $\{a_1, a_2, \cdots\} \subset (0,\infty)$ is *such that, for every nonnegative sequence* $\{b_1, b_2, \cdots\}$ *of real numbers such that,*

$\sum_{n=1}^{\infty} b_n^2 < \infty$, $\sum_{n=1}^{\infty} a_n b_n < \infty$, *then we have that,*

$$\sum_{n=1}^{\infty} a_n^2 < \infty$$

Clearly this result follows from corollary 2

The following theorem is known (cf. [1], p. 302).

LEMMA 1. Let (E,F,P) be a probability-space. Suppose $\{X_1, X_2, \cdots\}$ *is a sequence of mutually independent F-measurable real-valued functions and* $(\alpha_0, \alpha_1, \alpha_2, \cdots) \subset (0,\infty)$ *is such that,* $\sum_{n=1}^{\infty}$ *variance*(X_n) $= \infty$, *for all* $n = 1, 2, \cdots$, $|X_n| \le \alpha_n$ a.s. [P] and $\alpha_n^{-1}\sigma_n \to \alpha_0$ as $n \to \infty$, where, for all $n = 1,2,\ldots,$ $\sigma_n = \left[\sum_{j=1}^{n} \text{Variance}\,(X_j)\right]^{2^{-1}}$. Then for all $x \in (-\infty,\infty)$,

$$\lim_{n\to\infty} P\left(\omega \in E : \sigma_n^{-1}\left(\sum_{j=1}^{n} X_j(\omega)\right) \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$$

We now state our next theorem.

THEOREM 4. Let (E,F,P) be a probability-space. Suppose $(X_1, X_2, \ldots)$ is a sequence of mutually independent real-valued F-measurable functions such that, for each integer $n \geq 1$, X_n has Zero mean and finite variance σ_n^2. We assume that, there exists $k \in (0,\infty)$ such that, for all $n = 1, 2, \ldots\ldots,$

$$|X_n| \le K \left(\sum_{j=1}^{n} \sigma_j^2\right)^{2^{-1}} \quad \text{a.s. } [P]$$

Then the following statements are all equivalent :

(1) $$\limsup_{x\to\infty} \left[\limsup_{n\to\infty} P\left(\omega \in E : \sum_{j=1}^{n} X_j(\omega) \le x\right)\right] > 2^{-1}$$

(2) $\sum_{n=1}^{\infty}$ Variance $(X_n) < \infty$, and there exists a F-measurable real-valued random variable Y such that, $\left\{\sum_{j=1}^{n} X_j,\ n = 1,2,\cdots\right\}$ converges a.s.[p] to Y,

(3) $$\limsup_{n\to\infty} \sum_{j=1}^{n} X_j < \infty \quad \text{a.s. } [P]$$.

Proof. (1) Suppose equation (1) in the statement of theorem 4 holds. For each $n \in (1,2,\cdots)$, denote $\sigma_{1n}^2 = \sum_{j=1}^{n} \sigma_j^2$. We shall now show that,

(1.1) $$\lim_{n\to\infty} \sigma_{1n}^2 < \infty$$

To prove this, one can assume that, $\sigma_{1n}^2 > 0$, for all $n = 1,2,\ldots$. Suppose if possible, $\lim_{n\to\infty} \sigma_{1n}^2 = \infty$. Choose $x \in (0,\infty)$ and $\delta \in (0,\infty)$ such that,

(1.2) $$\limsup_{n\to\infty} P\left(\omega \in E : \sum_{j=1}^{n} X_j(\omega) \le x\right) > 2^{-1} + \delta$$

Choose $t \in (0,\infty)$ and $n_0 \in (1, 2, \cdots)$ such that, (1.3) $$(2\pi)^{-1/2}\int_{-\infty}^{t}\exp(-x'^2/2)\,dx' \le 2^{-1}+\frac{\delta}{4}$$,

and,

(1.4) for all $n \in (n_0, n_0+1, n_0+2, \cdots)$, $\sigma_{1n}^{-1}x \le t$

Then for every integer $n \geqslant n_0$,

$$P(\omega\in E:\sum_{j=1}^{n}X_j(\omega)\le x) = P(\omega\in E:\sigma_{1n}^{-1}\sum_{j=1}^{n}X_j(\omega)\le\sigma_{1n}^{-1}x)$$

$$\le P(\omega\in E:\sigma_{1n}^{-1}\sum_{j=1}^{n}X_j(\omega)\le t)$$ and so,

(1.5) $$\limsup_{n\to\infty} P(\omega\in E:\sum_{j=1}^{n}X_j(\omega)\le x) \le 2^{-1}+\frac{\delta}{4}$$

This contradicts equation (1.2) and this equation (1.1) holds. By a known martingale convergence theorem (cf.[7], p.137), equation (2) in the statement of theorem 3 holds.

(2) Equation (2) in the statement of theorem 4 trivially implies equation (3) in the statement of theorem 4. Suppose now, equation (3) in the statement of theorem 4 holds. Let $N\in F$ be such that, P(N) = 1 and for all $\omega\in N:\ \limsup_{n\to\infty}\sum_{j=1}^{n}X_j(\omega)<\infty$.

Then for all $\omega\in N$, there is an integer $K(\omega)\geqslant 1$ such that, for each $n\in(1,2,\cdots)$, $\sum_{j=1}^{n}X_j(\omega)\le K(\omega)$.

Using this, one can easily show that, there is an integer $K \geqslant 1$ such that,

$$P\left(\omega\in E:\sum_{j=1}^{n}X_j(\omega)\le K \text{ for all } n=1,2,\cdots\right) > \frac{3}{4}$$

One can easily show that, equation (1) in the statement of theorem 4 holds.

Theorem 5 which is being stated below follows from theorem 4.

THEOREM 5. Let (E,F,P) be a probability space. Suppose $(X_1, X_2, \cdots)$ is a sequence of mutually independent real-valued F-measurable functions such that, for each integer $n \geqslant 1$ X_n has Zero mean and $\sup_{n\geqslant 1}|X_n| \in L_1$ (E,F,P) (cf.[3], P.25). Then following statements are all equivalent:

(1) $$\limsup_{x\to\infty}\left[\limsup_{n\to\infty}P(\omega\in E:\sum_{j=1}^{n}X_j(\omega)\le x)\right] > 2^{-1}$$,

(2) there exists a F-measurable function $Y: E\to(-\infty,\infty)$ such that, $\{\sum_{j=1}^{n}X_j,\ n=1,2,\cdots\}$ converges $a.s.[P]$ to Y ,

(3) $$\limsup_{n\to\infty} \sum_{j=1}^{n} X_j < \infty \quad a.s. [P]$$

Proof. (1) Suppose equation (1) in the statement of theorem 4 holds. Let $\delta \in (0,\infty)$. Let $K \in (0,\infty)$ be such that, $\limsup_{n\to\infty} |X_n| < K$ a.s. $[P]$.

By Borel-Cantelli lemma (cf. [7], p.16), there exists $N \in F$ such that $P(N)=1$ and for all $\omega \in N$, $|X_j(\omega)| \leq K$, for all sufficiently large integral value of j. Using this and the hypothesis of the theorem, one can easily show that, there exists $A \in F$ such that $P(A) \geq 1-\delta$ and further the following conditions hold simultaneously :

(1.1) there exists an integer $m \geq 1$ such that, for each integer $n \geq m$ and for all $\omega \in A$,

$$X_n(\omega) = X_n(\omega) I_{A_{n,K}}(\omega) \quad ,$$

where for each n = 1,2, ..., $A_{n,K} = (\omega \in E : |X_n(\omega)| \leq K)$,

and for every $A \in F$, I_A denotes indicator function of theset A ,

(1.2) there exists $T \in (0,\infty)$ such that, for all $\omega \in A$ and for all i = 1, ... m, $|X_j(\omega)| \leq T$

By ([6], P.151) and the proof of theorem 3.1 in ([5], p.55),

(1.3) $$\sum_{n=1}^{\infty} \int_{E-A_{n,K}} |X_n| \, dP < \infty$$

Since mean of each X_n is 0, by equation (1.3),

(1.4) $$\sum_{n=1}^{\infty} \left| \int_{A_{n,K}} X_n \, dP \right| < \infty$$

By the hypothesis of the theorem and equations (1.1), (1.2) and (1.4), we get that, there exists $x \in (0,\infty)$ such that,

(1.5) $$\limsup_{n\to\infty} P\Big(\omega \in E : \sum_{j=1}^{n} \big(X_j(\omega) I_{A_{j,K}}(\omega) - E(X_j I_{A_{j,K}})\big) \leq x\Big) > 2^{-1}.$$

By theorem 3 and by equations (1.3), (1.4) and monotone convergence theorem (cf.[7],p.39), equation (2) in the statement of theorem 4 holds.

(2) The fact that equation (2) in the statement of theorem 4 implies equation (3) in the statement of theorem 4 is trivial. The proof of the fact that, equation (3) in the statement of theorem 4 implies equation (1) in the statement of theorem 4 uses an argument that has already been made (cf. proof of theorem 3).

Equivalence of equations (2) and (3) in the statement of theorem 3 is known (cf. [7], p.150).

The following result emerges from theorem 4.

COROLLARY 3. Let (E,F,P) be a probability space. Suppose $(X_1, X_2, \cdots)$ is a sequence of mutually independent real-valued F-measurable functions such that, there are real numbers $K \in (0, \infty)$, $m \in (-\infty, \infty)$ such that, each X_n has mean m, for each $n \in (1, 2, \cdots)$,

$$\text{Variance}\,(X_n) \leq K, \qquad |X_n| \leq K\left(\sum_{j=1}^{n} \text{Variance}\,(X_j)\right)^{2^{-1}}$$

a.s. (P) and

$$\sum_{n=1}^{\infty} \text{Variance}\,(X_n) = \infty .$$

Then for any two real numbers a and b such that $a < b$,

$$P\left(\omega \in E : a \leq \sum_{j=1}^{n} X_j(\omega) \leq b \quad \text{for all } n \geq 1\right) = 0$$

Proof. Suppose if possible, there are real numbers a and b such that, $a < b$ and,

(1) $$P(A) > 0$$

where,

(2) $$A = \left(\omega \in E : a \leq \sum_{j=1}^{n} X_j(\omega) \leq b \quad \text{for all } n \geq 1\right)$$

For all $\omega \in A$, $n^{-1}\sum_{j=1}^{n} X_j(\omega) \to 0$ as $n \to \infty$. By the strong law of large numbers (cf.[7], P. 148), $\{n^{-1}\sum_{j=1}^{n}(X_j - m),\ n = 1, 2, \cdots\}$ converges a.s. [P] to 0 and hence, m = 0. Since $P(A) > 0$,

$\limsup_{n \to \infty} \sum_{j=1}^{n} X_j < \infty$ a.s. [P] and consequently, by theorem 4, $\sum_{n=1}^{\infty}$Variance $(X_n) < \infty$. Thus, P (A) = 0.

The following known result (Cf.[10]) emerges as a corollary to corollary 3.

COROLLARY 4. Let (E,F,P) be a probability space. Suppose $(X_1, X_2, \cdots)$ is sequence of mutually independent, identically distributed real-valued F-measurable functions such that, $P(\omega \in E : X_1(\omega) \neq 0) > 0$. Then for any two real numbers a and b such that, $a < b$,

$$P(\omega \in E : a \leq \sum_{j=1}^{n} X_j(\omega) \leq b, \text{ for all } n \geq 1) = 0$$

Proof. Suppose if possible, there are real numbers a and b such that, $a < b$ and

$$P(\omega \in E, a \leq \sum_{j=1}^{n} X_j(\omega) \leq b, \text{ for all } n \geq 1) > 0.$$

Let $A = (\omega \in E : a \leq \sum_{j=1}^{n} X_j(\omega) \leq b$, for all $n \geq 1)$. $P(A) > 0$.

There is a real number $K \in (0, \infty)$ such that,

$P(\omega \in E : 0 < |X_1(\omega)| \leq K) > 0$ and, $A \subset [\omega \in E : |X_j(\omega)| \leq K$,

for each integer $j \geq 1]$. Then

(2) $$P(B) > 0$$

where,

(3) $$B = [\omega \in E : a \leq \sum_{j=1}^{n} X_j(\omega) \, I_{B_j}(\omega) \leq b, \text{ for all } n \geq 1],$$

and for all $j = 1, 2, \cdots$,

(4) $$B_j = [\omega \in E : |X_j(\omega)| \leq K]$$

Let $Y_j(\omega) = X_j(\omega) \, I_{B_j}(\omega)$, for all $j = 1, 2, \cdots$. Each Y_j has the same mean m and variance σ^2. $\sigma^2 > 0$ By corollary 3, corollary 4 follows.

REFERENCES

1. CHOW, YUAN. SHIH AND TEICHER, HENRY, (1979) : "Probability theory, Independence, Interchangeability and Martingales", Narosa Publishing House, New Delhi.

2. CONWAY, JOHN. B. (1980):" Functions of one complex variable", second edition, Springer Verlag, New York.

3. GHOSH, S. (1991) : " Some results related to characterisation of distribution functions and some applications", Proceedings of the international A.M.S.E. Conference held in New Delhi, India during December 9-11, 1991, p.19-31, Volume I, A.M.S.E. Press, France.

4. GHOSH, S. (1991) : "Applications of methods of numerical analysis in solving ring-valued equations and an integral equation leading to a characterisation of the rectangular distribution", Proceedings Intern. A.M.S.E. Delhi conference, (1991), Vol. I, P. 33-45.

5. GHOSH, S. (1991) : "On some theorems on perfect Probabilities and that strong law of large numbers", proceedings of the international A.M.S.E. Conference held in New-Delhi, India during December 9-11, 1991, p.47-59, VOLUME I, A.M.S.E. Press, France.

6. LOEVE, M. (1963) : "Probability theory, third edition, Van Nostrand, Princeton, New Jersey.

7. NEVEU, J. (1965) : "Mathematical Foundations of the Calculus of Probability, Holden Day, San Fransisco.

8. RUDIN, W. (1974) : "Real and complex analysis", McGraw Hill, New York.

9. SIMMONS, G.F. (1963) : "Introduction to Topology and modern analysis", McGraw Hill, New York.

10. STEIN, C. (1946) : "A note on cumulative sums", Ann. Math. Stat. 17, p.498-499.

11. USPENSKY, J. V. (1965) : "Introduction to Mathematical Probability, McGraw Hill, New York.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 53-66

A PROBABILISTIC APPROACH TO STONE-WEIERSTRASS APPROXIMATION THEOREM

BY

S. GHOSH

Department of Statistics, Presidency College, Calcutta-700 073

INDIA

1. ABSTRACT AND INTRODUCTION.

Let (Ω, τ) be a compact Hausdorff topological space, let $C(\Omega, R) = \{f : f : \Omega \to R$ is a continuous function$\}$. Stone-Weierstrass approximation theorem gives a sufficient condition under which a closed vector subspace of $C(\Omega, R)$ is equal to $C(\Omega, R)$. In this paper, using a probabilistic approach, we have generalised the above theorem by taking instead of an element of $C(\Omega, R)$, a continuous function $f : \Omega \to V$, where V is a topological vector-space. Our results seem to be new.

2. DEFINITIONS : NOTATIONS : TERMINOLOGIES :

Throughout our subsequent discussions, (Ω, τ) will denote an arbitrary topological space. Denote

(2.1) $M_1(\Omega, R) = \{f : f : \Omega \to (-\infty, \infty)$ is a continuous function such that, $\sup_{x \in \Omega} |f(x)| < \infty\}$,

(2.2) $M_2(\Omega, R) = \{f : f : \Omega \to (-\infty, \infty)$ is a function such that, $\sup_{x \in \Omega} |f(x)| < \infty\}$,

(2.3) $N = \{0, 1, -1, 2, -2, \cdots\}$, $R = (-\infty, \infty)$, $\phi =$ empty set

Let M be a nonempty subset of $M_2(\Omega, R)$. An element $f \in M_2(\Omega, R)$ is called a M-finite-dimensional trigonometric polynomial, if there exists a *nonempty finite subset* $(g_1, \cdots g_t)$ *of* M *and an integer* $K \geq 1$ *such that, for each* $j \in (1, \cdots K)$, *there is a real number* α_j *and a sequence* $(\alpha_{j1}, \cdots \alpha_{jt}) \subset N$ *such that, for all* $x \in \Omega$,

(2.4) $$f(x) = \sum_{j=1}^{K} \alpha_j \left[\cos\left(\sum_{m=1}^{t} \alpha_{jm} g_m(x)\right) + \sin\left(\sum_{m=1}^{t} \alpha_{jm} g_m(x)\right)\right]$$

A function $f \in M_1(\Omega, R)$ *is called* (Ω, τ, M)*- trigonometrically approximable, if for each* $\delta \in (0, \infty)$, *there is an element* $g_\delta \in M_2(\Omega, R)$ *which is a* M *-finite-dimensional trigonometric polynomial and which is such that,* $\sup_{x \in \Omega} |f(x) - g_\delta(x)| \leq \delta.$ *Let* α *be an ordinal number such that,* $\alpha \geq 1$. *Let* A *be a nonempty set. Define*

(2.5) $$(A)^\alpha = \left((x_\beta : \beta \text{ ordinal}, \beta < \alpha) : \text{each } x_\beta \in A \right)$$

Suppose $\Phi \neq A \subset (R)^\alpha$. *A function* $f : A \to (-\infty, \infty)$ *will be called a* A *-semi-finite dimensional trigonometric polynomial, if there are strictly positive integers* K, t *and sequences* $(a_1, \cdots a_K) \subset (-\infty, \infty)$, $(\beta_1, \cdots \beta_t) \subset (\beta : \beta \text{ ordinal}, \beta < \alpha)$ *such that, for all* $j = 1, \cdots K$, *there is a sequence* $(\alpha_{j1}, \cdots, \alpha_{jt}) \subset N$ *such that, for all* $(x_\beta : \beta \text{ ordinal and } \beta < \alpha) \in A$,

(2.6) $$f(x_\beta : \beta \text{ ordinal and } \beta < \alpha) = \sum_{j=1}^{K} a_j \left[\cos\left(\sum_{m=1}^{t} \alpha_{jm} x_{\beta_m}\right) + \sin\left(\sum_{m=1}^{t} \alpha_{jm} x_{\beta_m}\right)\right]$$

Let $f : A \to (-\infty, \infty)$ *be a function.* f *is said to be trigonometrically approximable on* A, *if for each* $\delta \in (0, \infty)$, *there is a* A *-semi-finite-dimensional trigonometric polynomial* $g_{\delta, f} : A \to (-\infty, \infty)$ *such that, for all* $x \in A$, $|f(x) - g_{\delta,f}(x)| \leq \delta$. *Let* σ *denote Borel-*σ*-algebra on* $(R)^\alpha$ *relativised to* A. *Let* $m : \sigma \to [0, \infty]$ *be a sigma-finite measure. Let* $p \in [1, \infty), f \in L_p(A, \sigma, m), g \in L_p(A, \sigma, m)$ *(cf. [2], p.25).* f *is said to be* (p, σ, m, g) *-trigonometrically approxi-*

mable, if for each $\delta \in (0,\infty)$, there is a A -semi-finite-dimensional trigonometric polynomial $h_{g,\delta} : A \to (-\infty, \infty)$ such that,

$$\int_A \left| f(x) - h_{g,\delta}(x)\chi(x) \right| m(dx) \leq \delta$$

Let E be a nonempty subset of $M_2(\Omega, R)$. E is called an algebra if $g_1 \in E$, $g_2 \in E$ always implies that, $g_1 g_2 \in E$. E is called a semi-algebra if given strictly positive integers K, t, a finite sequence $(g_1, \cdots\cdots g_t) \subset E$, and sequences $(\alpha_1, \cdots \alpha_K) \subset (R)^K$ and $(\alpha_{j1}, \cdots \alpha_{jt}, j = 1, \cdots K) \subset (N)^{Kt}$

the function f defined by,

$$f(x) = \sum_{j=1}^{K} \alpha_j \left[\cos\left(\sum_{m=1}^{t} \alpha_{jm} g_m(x) \right) + \sin\left(\sum_{m=1}^{t} \alpha_{jm} g_m(x) \right) \right]$$

for all $x \in \Omega$, has the property that, $f \in E$. E is said to seperate points of Ω, if given distinct elements x and y of Ω, there exists an element $g_{x,y}$ of E such that, $g_{x,y}(x) \neq g_{x,y}(y)$. Let V be a topological vector space whose zero element is denoted by θ. Let $f : \Omega \to V$ be a function. f is called continuous, if for each open subset V_1 of V, $f^{-1}(V_1) \in \mathcal{T}$. Let $g : \Omega \to (-\infty, \infty)$ be a function. Let $(n_1, n_2, \cdots)$ be a sequence of distinct strictly positive real numbers. f is called a $(g, (n_1, n_2, \cdots), V)$ - polynomial, if there is a strictly positive integer K and $(a_0, \cdots a_K) \subset V$ such that, $\forall\ x \in \Omega$,

$$f(x) = a_0 + \sum_{j=1}^{K} (g(x))^{n_j} a_j$$

A function $f_1 : \Omega \to V$ is called a (g, V) -cosine-polynomial if there is an integer $K \geq 1$ and $(a_0, \cdots\cdots a_K) \subset V$ such that, for all $x \in \Omega$,

$$f_1(x) = a_0 + \sum_{j=1}^{K} \cos(j\, g(x))\, a_j .$$

f is said to be uniformly approximable by g, if given any open subset V_1 of V such that, $\theta \in V_1$ and a sequence $(n_1 < n_2 < \cdots\cdots) \subset (0, \infty)$ such that,

$\sum_{j=1}^{\infty} (n_j)^{-1} = \infty$, there is a $(g, (n_1, n_2, \cdots), V)$-polynomial h and a (g, V)-cosine-polynomial h_1 such that, for all $x \in \Omega$, all the four elements $(f(x) - h(x))$, $(h(x) - f(x))$, $(f(x) - h_1(x))$ and $(h_1(x) - f(x))$ are elements of V_1. Let α be an ordinal $\geqslant 1$. Let $\phi \neq D \subset (R)^{\alpha}$. Let $S_{1,\alpha,D} =$

$[A : A = D \cap (\prod_{\beta : \beta \text{ ordinal}, \beta < \alpha} A_\beta)$, where there is a finite sequence $(\beta_1, \cdots\cdots \beta_K) \subset (\beta : \beta$ ordinal, $\beta < \alpha)$ such that, if β be an ordinal $< \alpha$, $A_\beta = R$, if $\beta \notin (\beta_1, \cdots\cdots \beta_K)$ whereas if $\beta \in (\beta_1, \cdots\cdots \beta_K)$, A_β isan open subset of $R]$.

We shall denote,

(2.7) $\tau_{\alpha,D} = (A : A$ is expressible as union of some of the elements of $S_{1,\alpha,D}$. We observe that $\tau_{\alpha,D}$ is a topology on D.

3. THE MAIN RESULTS.

THEOREM 1. Let α be an ordinal $\geqslant 1$. Let $\alpha_1 = ((-\pi, \pi))^{\alpha}$, let σ denote Borel-σ-algebra on α_1, let $\mu : \sigma \to [0, \infty]$ be a sigma-finite-measure. Let D be a nonempty compact subset of $(R)^{\alpha}$ such that, $D \subset \alpha_1$. Let $p \in [1, \infty)$. Let $h \in L_p(\alpha_1, \sigma, \mu)$ be such that, $\mu(x \in \alpha_1 : h(x) = 0) = 0$. Let $f : D \to (-\infty, \infty)$ be a continuous function, let $f_1 \in L_p(\alpha_1, \sigma, \mu)$ (cf.[2], p.25). Then,

(1.1) f is trigonometrically approximable on D

(1.2) f_1 is (p, σ, μ, h)-trigonometrically approximable.

Proof. (1) Let C denote the class of all continuous functions defined on D. For all $f \in C$, define $|f|_D = \sup_{x \in D} |f(x)|$. Let

(1.1) $T_1 = \{ f' \in C : f'$ is trigonometrically approximable on D $\}$,

(1.2) $T_2 = \{ f' \in L_p(\alpha_1, \sigma, \mu) : f'$ is (p, g, μ, g)-trigonometrically approximable $\}$.

For each $j \in \{1, 2\}$, $T_j \neq \phi$. T_1 is a closed linear subspace of C,

T_2 is a closed linear subspace of $L_p(\alpha_1, \sigma, \mu)$. We shall first of all show that,

(1.3)
$$T_1 = C$$

Let $K \in (1, 2, \dots)$, $(a_1, \dots a_K) \subset N$, $(\beta_1, \dots \beta_K) \subset (\beta : \beta$ ordinal, $\beta < \alpha)$. Let $x = (x_\beta : \beta$ ordinal, $\beta < \alpha) \in \alpha_1$ Define

(1.4)
$$h_{(1, a_1, \dots a_K, \beta_1, \dots \beta_K)}(x) = \cos\left(\sum_{j=1}^{K} a_j x_{\beta_j}\right)$$

(1.5)
$$h_{(2, a_1, \dots a_K, \beta_1, \dots \beta_K)}(x) = \sin\left(\sum_{j=1}^{K} a_j x_{\beta_j}\right)$$

Let $(l_{j1}, l_{j2}, j = 1, \dots K) \subset R$. Define

(1.6)
$$h_{(3, l_{j1}, l_{j2}, \beta_j, j = 1, \dots K)}(x)$$

$$= \exp\left[i \sum_{j=1}^{K} l_{j1} \cos x_{\beta_j} + i \sum_{j=1}^{K} l_{j2} \sin x_{\beta_j}\right],$$ where $i = \sqrt{-1}$.

(2) Suppose if possible, $T_1 \neq C$. Then there exists $f \in C$ such that, $f \notin T_1$. Using some properties of trigonometric functions and some known theorems (cf. [3], p.247,249, [5], p.42,49 and [6], p. 229), one can easily show that, there exists a measure $\mu_1 : \sigma \cap D \to (-\infty, \infty)$ such that,

(2.1)
$$\int_D f(x) \mu_1(dx) = 1$$ but, for each $K \in (1, 2, \dots)$,

$(a_1, \dots a_K) \subset N$, $(\beta_1 \dots, \beta_K) \subset (\beta : \beta$ ordinal, $\beta < \alpha)$,

$$\left| \int_D h_{(1, a_1, \dots a_K, \beta_1, \dots \beta_K)}(x) \mu_1(dx) \right| +$$
$$\left| \int_D h_{(2, a_1, \dots a_K, \beta_1, \dots \beta_K)}(x) \mu_1(dx) \right| = 0$$

Let $A \in \sigma$. Let $\mu_1^+(A) = \text{Sup}\{\mu_1(B) : B \subset A, B \in \sigma\}$,

$\mu_1^-(A) = \{-\mu_1(B) : B \subset A, B \in \sigma\}$. From (2.1) and a known result (cf.[4], p.105),

(2.2)
$$\mu_1^+(D) = \mu_1^-(D) = A$$ (Say)

$A \in (0, \infty)$, because if $A = 0$, (2.1) get contradicted. Using (2.1), (2.2) and some known properties of trigonometric functions, one can easily show that, for each $K \in (1.2, \ldots\ldots)$, for each sequence $(l_{j1}, l_{j2}, j = 1, \cdots K) \subset R$, and for each $(\beta_1, \cdots \beta_K) \subset (\beta : \beta$ ordinal, $\beta < \alpha)$,

$$(2.3) \qquad A^{-1} \int_D h_{(3, l_{j1}, l_{j2}, \beta_j, j=1, \cdots K)}(x) \quad \mu_1^+ (dx)$$

$$= A^{-1} \int_D h_{(3, l_{j1}, l_{j2}, \beta_j, j=1, \cdots K)}(x) \; \mu_1^- (dx)$$

By a known theorem (cf.[1], p.101), for each $K \in (1, 2, \cdots\cdots)$, $(\beta_1, \cdots \beta_K) \subset (\beta : \beta$ ordinal , $\beta < \alpha$) and for each Borel Subset B of the 2K-dimensional Eucledian-space ,

$$(2.4) \qquad \mu_1^+ \Big((x_\beta : \beta \text{ ordinal}, \beta < \alpha) \in D :$$

$$(\cos x_{\beta_1}, \sin x_{\beta_1}, \cdots \cos x_{\beta_K}, \sin x_{\beta_K}) \in B \Big)$$

$$= \mu_1^- \Big((x_\beta : \beta \text{ ordinal}, \beta < \alpha) \in D :$$

$$(\cos x_{\beta_1}, \sin x_{\beta_1}, \cdots \cos x_{\beta_K}, \sin x_{\beta_K}) \in B \Big)$$

Using this one can easily show that, for each Borel subset $B' \bullet f(R)^{\alpha}$, $\mu_1^+ \big((x_\beta : \beta$ ordinal , $\beta < \alpha)$:

$$(\cos x_\beta \;, \sin x_\beta \;:\; \beta \text{ ordinal}, \beta < \alpha) \in B' \big)$$

$$\mu_1^- = \big((x_\beta \;: \beta \text{ ordinal}, \beta < \alpha) : (\cos x_\beta, \sin x_\beta : \beta$$

$$\text{ordinal}, \beta < \alpha) \in B' \big)$$

. Define a function $B : D \to (R)^{\alpha}$ as follows : Let $(x_\beta : \beta$ ordinal , $\beta < \alpha) \in D$. Then

$$B(x_\beta : \beta \text{ ordinal}, \beta < \alpha) = (\cos x_\beta \;, \sin x_\beta \;:\; \beta \text{ ordinal}, \beta < \alpha).$$

If $A \in \tau_{\alpha, (R)^{\alpha}}$ (cf. equation (2.7) in section 2), $[x \in D : B(x) \in A]$ is expressible as $D \cap \mathcal{U}$, where $\mathcal{U} \in \tau_{\alpha, (R)^{\alpha}}$. Using some known theorems (cf. [6], p.111 and 131) and the fact that, B is an one-to-one *continuous function, one can show that, for every Borel subset* B' *of* D, $B(B')$ *is a Borel subset of* $(R)^{\alpha}$. *Using this and some of the facts stated earlier, one can show that, for every* $B' \in \sigma \cap D$, $\mu_1(B') = 0$. *This contradicts equation (2.1) and thus,* $T_1 = C$.

(3) We shall now prove that,

(3.1)
$$T_2 = L_p(\alpha_1, \sigma, \mu)$$

Suppose if possible equation (3.1) is false. Let q *be such that,* $p^{-1} + q^{-1} = 1$, *where if P = 1,* q *should be interpreted to be* ∞. *Using the proof of theorem 3.2 in ([2], p.25) and some properties of the trigonometric functions, one can easily show that, there exists* $f \in L_p(\alpha_1, \sigma, \mu)$, $g \in L_q(\alpha_1, \sigma, \mu)$ *such that,*

(3.2) $\int_{\alpha_1} f(x)\, g(x)\, \mu(dx) = 1$ *but for every integer* $K \geq 1$, *for every sequence* $(a_1, \dots a_K) \subset N$, *for every* $(\beta_1, \dots \beta_K) \subset (\beta : \beta$ *ordinal*, $\beta < \alpha)$, *and for every integer* $j \in \{1, 2\}$ *(cf.equations (1.4) and (1.5)),* $\int_{\alpha_1} h_{(j, a_1, \dots a_K, \beta_1, \dots \beta_K)}(x)\, g(x)\, h(x)\, \mu(dx) = 0$

Using equation (3.2) and a slight modification of the argument thathas already been used, one can show that,

(3.3)
$$\mu(x \in \alpha_1 : g(x) \neq 0) = 0$$

Using (3.3) and the hypothesis of the theorem, one can show that, $\mu(x \in \alpha_1 : g(x) \neq 0) = 0$ *and this proves theorem 1.*

Theorem 1 generalises the known result (cf. [5], p.34) that, if $f : (-\pi, \pi) \to (-\infty, \infty)$ *be a Borel-measurable square integrable (with respect to Lebesgue-measure) function, then there is an integer n* $\geq$ *1 and sequences*

$$(K_{11}, \dots K_{1n}, K_{21}, \dots K_{2n}) \subset (N)^{2n}$$

$$(a_{1n} \dots, a_{1n}, a_{21} \dots, a_{2n}) \subset (R)^{2n}$$ *such that,*

$$\int_{-\pi < x < \pi} \Big[f(x) - \sum_{j=1}^{n} \{ a_{1j} \cos(K_{1j} x) + a_{2j} \sin(K_{2j} x) \} \Big]^2 dx \leq \delta$$

THEOREM 2. Let $M_3(\Omega, R) = \{ f : f : \Omega \to [0, 2^{-1}]$ is a function $\}$. Suppose α is an ordinal ≥ 1 and suppose $M' = (f_\beta : \beta$ Ordinal, $\beta < \alpha) \subset M_3(\Omega, R)$ is such that, the function $g : \Omega \to (R)^\alpha$ defined by,

$$g(x) = (f_\beta(x) : \beta \text{ ordinal}, \beta < \alpha),$$

for all $x \in \Omega$, has the properties that, g is one-to-one, $g(\Omega)$ is a compact subset of $(R)^\alpha$, and further, the function g is an open mapping, in the sense that, for each $A \in \tau$, $g(A) \in \tau_{\alpha, g(\Omega)}$ (cf. equation (2.7) in section 2).

Then each $f \in M_1(\Omega, R)$ is (Ω, τ, M')-trigonometrically approximable. Further if $\alpha \leq$ ordinal number of $\{0, 1, 2, \cdots\}$, there is an one-to-one continuous function $g' : \Omega \to (0,1)$ such that, given any topological vector space V and a continuous function $f' : \Omega \to V$, f' is uniformly approximable by g'.

Proof. (1) For each $x \in g(\Omega)$, let $g^{-1}(x)$ denote the unique $y \in x$ such that, $g(y) = x$. Let $f \in M_1(\Omega, R)$. For each $x \in g(\Omega)$, define $f'(x) = f(g^{-1}(x))$. Using the hypothesis of the theorem, one can show that, $f' : g(\Omega) \to (-\infty, \infty)$ is a continuous function. By theorem 1, f' is trigonometrically approximable on $g(\Omega)$. Using this, one can easily show that,

f is (Ω, τ, M')-trigonometrically approximable.

(2) Let

(2.1) ω = ordinal number of $\{ 0, 1, 2, \ldots\ldots \}$

Suppose $\alpha \leq \omega$. consider the case, $\alpha < \omega$. There is an integer $n \geq 0$ such that $\alpha = n+1$. For every integer $m > n$, define $f_m(x) = 0$, for all $x \in \Omega$. Using this modification, one can easily assume without loss of generality that, $\alpha = \omega$. Let $(y_1, y_2, \cdots) \subset ([0, 1))^\omega$. Let $j \in (1, 2, \cdots)$. Write $y_j = \sum_{n=1}^{\infty} 2^{-n} y(j, n)$, where each $y(j,n) \in \{0, 1\}$, and there does not exist an integer $n \geq 1$ such that, for each integer $m \geq n$, $y(j, m) = 1$ Let

(2.2) $f_1'(y_1, y_2, \cdots) = \frac{1}{2} y(1,1) + \frac{1}{2^2} y(2,1) + \frac{1}{2^3} y(1,2) + \frac{1}{2^4} y(3,1) + \frac{1}{2^5} y(2,2) + \frac{1}{2^6} y(1,3) + \cdots$

Suppose $\{(y_1, y_2, \cdots), (x_{j1}, x_{j2}, \cdots), j = 1, 2, \cdots\} \subset ([0,1))^{\omega}$ and further, $\sum_{k=1}^{\infty} 2^{-k} |x_{jk} - y_k| \to 0$ as $j \to \infty$. Then $f_1'(x_{j1}, x_{j2}, \cdots) \to f_1'(y_1, y_2, \cdots)$ as $j \to \infty$. $f_1' : ([0,1))^{\omega} \to [0,1]$ is an one-to-one continuous function. For all $x \in \Omega$, define

(2.3) $$h_1(x) = 3^{-1}(f_1'(g(x)) + 1)$$

$h_1(\Omega) \subset (0,1)$ and further, $h_1(\Omega)$ is a compact subset of the real line. One can easily show that, for every set $A \in \tau$, $h_1(A) = A' \cap h_1(\Omega)$, where A' is an open subset of R. Let $f \in M_1(\Omega, R)$. Let $x \in h_1(\Omega)$. Let $h_1^{-1}(x)$ denote the unique $y \in \Omega$ such that, $h_1(y) = x$. Define $g_1(x) = f(h_1^{-1}(x))$. $g_1 : h_1(\Omega) \to (-\infty, \infty)$ is a continuous function. By theorem 1, g_1 is trigonometrically approximable on $h_1(\Omega)$. Both $\sin x$ and $\cos x$ admit of power series expansions. Let $\delta \in (0,\infty)$. Using these facts, one can show that, there is an integer $n \geqslant 1$ and a sequence $\{a(j, \delta), j = 0, 1, \cdots n\} \subset (-\infty, \infty)$ such that,

(2.4) $$\sup_{x \in h_1(\Omega)} \left| g_1(x) - a(0, \delta) - \sum_{j=1}^{n} a(j, \delta) x^j \right| \leq \frac{\delta}{4}$$

Let $\delta_1 \in (0, \infty)$. Let $x \in h_1(\Omega)$. Suppose $x \neq 0$

$$\lim_{\substack{n \to \infty \\ n \in (1, 2, \cdots)}} \left(\frac{x}{n}\right)^{-1} \sin\left(\frac{x}{n}\right) = 1$$

There exists $N(x) \in (1, 2, \cdots)$ such that, for each integer $n \geqslant N(x)$,

$$\left| \left(\frac{x}{n}\right)^{-1} \sin\left(\frac{x}{n}\right) - 1 \right| \leq \left[4\left(\sup_{x \in h_1(\Omega)} |x| + 1\right)\right]^{-1} \delta_1$$

Thus for each integer $n \geqslant N(x)$, $\left| n \sin\left(\frac{x}{n}\right) - x \right| \leq \frac{\delta_1}{4}$

This is true even when $x = 0$. By first mean value theorem of differential calculus, for every pair of real numbers x and y and for all $n \in (1, 2, \cdots)$, $\left| n \sin\left(\frac{x}{n}\right) - n \sin\left(\frac{y}{n}\right) \right| \leq |x - y|$. So given $x \in h_1(\Omega)$, there exists $\delta_1(x) \in (0, \infty)$ such that, whenever

$y \in h_1(\Omega) \cap (x-\delta_1(x), x+\delta_1(x))$ and $n \in (N(x), N(x)+1, \cdots\cdots)$,

$|n\sin(\frac{y}{n}) - y| \le |n\sin(\frac{y}{n}) - n\sin(\frac{x}{n}) + n\sin(\frac{x}{n}) - x$

$+ x - y| \le \delta_1$. Using an argument involving Heine-Borel-theorem (cf.[6],p.119), one can easily show that,

(2.5) there is an integer $M(\delta_1) \geq 1$ such that, for each integer $m \geq M(\delta_1)$ and for each $y \in h_1(\Omega)$,

$$|m\sin(\frac{y}{m}) - y| \le \delta_1 \quad .$$

Using (2.4) and (2.5), one can easily show that, there is an integer t $\geq$ 1 and a sequence $\{a'(j,\delta), j=0,1,\cdots\cdots\cdots n\} \subset (-\infty, \infty)$ such that,

$$(2.6)\quad \sup_{x \in h_1(\Omega)} |g_1(x) - a'(0,\delta) - \sum_{j=1}^{n} a'(j,\delta)\left(\sin\left(\frac{x}{2^t}\right)\right)^j| \le \frac{\delta}{2} \quad .$$

For all m = 1, ... t, $\sin(2^{-m}x)$ is expressible as a power series in $\cos(2^{-(m-1)}x)$ and $\cos(2^{-m}x)$ is expressible as a power series in $\sin(2^{-m}x)$. Using these facts, one can easily show that, there is an inter k $\geq$ 1 and a sequence $\{b(j,\delta), j=0,1,\cdots K\} \subset (-\infty,\infty)$ such that,

$$(2.7)\quad \sup_{x \in h_1(\Omega)} |g_1(x) - b(0,\delta) - \sum_{j=1}^{K} b(j,\delta)\cos(jx)| \le \delta \quad .$$

(3) Let V be a topological vector-space whose Zero element is denoted by θ. Let $f : \Omega \to V$ be a continuous function. To show that, f is uniformly approximable by h_1, we assume without loss of generality that, Ω is not a singleton set. Let $a_1 =$ Inf $\{x : x \in h_1(\Omega)\}$,

$$a_2 = \sup\{x : x \in h_1(\Omega)\}.$$

$a_1 < a_2$. $[a_1, a_2] \subset (0,1)$. Define $f_1(x) = f(h_1^{-1}(x))$, $x \in h_1(\Omega)$. Let V_1 be an arbitrary open subset of V such that, $\theta \in V_1$. Let V_2, V_3, V_4 be open subsets of V such that, for all $j \in \{2,3,4\}$, $\theta \in V_j$ and the following conditions hold (cf. [2], p. 29):

(3.1) $(x_1, x_2) \in V_2 \times V_3$ always implies that, $(x_1 + x_2) \in V_1$,

(3.2) $(x_1, x_2) \in V_4 \times V_4$ always implies that, $(x_1 + x_2) \in V_2$, (3.3) there exists $\delta \in (0,\infty)$ such that, whenever $x \in h_1(\Omega)$, $y \in h_1(\Omega)$ and $|x-y| \le \delta$, for all $\lambda \in [-1,1]$,

$$\lambda(f(x)-f(y)) \in V_4$$

Choose a finite sequence of points $(x_0 = a_1 < x_1 < \cdots < x_n = a_2) \subset h_1(\Omega)$ such that, for each integer j = 0,1, ... n-1, either $(x_{j+1} - x_j) \le \delta$ or $(x_j, x_{j+1}) \cap h_1(\Omega) = \phi$. Let $j \in (0,1,\ldots n-1)$. For all $y \in B_j$, define $f_1(y) = f(x_j) + (y - x_j)(x_{j+1} - x_j)^{-1}(f(x_{j+1}) - f(x_j))$, where if $j < n-1$, $B_j = [x_j, x_{j+1})$, whereas if $j = n-1$, $B_j = [x_{n-1}, a_2]$.

For all $y \in h_1(\Omega)$, both $(f(y) - f_1(y))$ and $(f_1(y) - f(y))$ are elements of V_2 (cf. [2], p.29). As in the proof of theorem 3.3 in ([2],p.29), one can show that, there is a sequence $(b_0, \cdots b_{n-1}) \subset V$ such that,

(3.4) for all $x \in [a_1, a_2]$, $f_1(x) - f(a_1) = \sum_{j=0}^{n-1} \left[\int_{a_1 \le t \le x} I_{B_j}(t)\,dt \right] b_j$,

where for every Borel subset B of $[a_1, a_2]$, I_B denotes indicator function of the set B. There is a sequence $(\lambda_0, \cdots \lambda_{n-1}) \subset (0,1)$ such that, whenever $(\alpha'_0, \cdots \alpha'_{n-1})$ are real numbers such that, $|\alpha'_j| \le \lambda_j$, for all $j = 0, \cdots n-1$, $\sum_{j=0}^{n-1} \alpha'_j b_j \in V_3$. Let $(n_1 < n_2 < \cdots) \subset (0,\infty)$ be such that, $\sum_{j=1}^{\infty} (n_j)^{-1} = \infty$. Since $[a_1, a_2] \subset (0,1)$, by equation (2.7) and a known theorem (cf. [5], p. 337), there is an integer $k \geqslant 1$ and a sequence $(a_m(0,j), \cdots a_m(K,j), m \in \{1,2\}, j = 0, \cdots n-1) \subset R$ such that, for all $x \in [a_1, a_2]$ and for all $j = 0, \cdots n-1$,

(3.5) $$\left| \int_{a_1 \le t \le x} I_{B_j}(t)\,dt - a_1(0,j) - \sum_{t=1}^{K} a_1(t,j)\, x^{n_t} \right| \le \lambda_j$$

(3.6) $$\left| \int_{a_1 \le t \le x} I_{B_j}(t)\,dt - a_2(0,j) - \sum_{t=1}^{K} a_2(t,j) \cos(t x) \right| \le \lambda_j .$$

Using equations (3.4), (3.5) and (3.6), one can easily prove the truth of last part of theorem 2.

Throughout our subsequent discussions, (Ω', τ') will denote a Hausdorff compact topological space, $C(\Omega', R)$ will denote the class of real-valued continuous functions defined on Ω'. If $f \in C(\Omega', R)$, $\sup_{x \in \Omega'} |f(x)| < \infty$.

If $f \in C(\Omega', R)$, $g \in C(\Omega', R)$, define $d(f, g) = \sup_{x \in \Omega} |f(x) - g(x)|$.

In our subsequent discussions, E will denote a nonempty closed vector subspace of $C(\Omega', R)$ which contains constant functions and which is a semi-algebra.

THEOREM 3. We assume that,

(1) E seperates points of Ω'

Then,

(2) $$E = C(\Omega', R)$$

Further if (Ω', τ') be a second-countable topological space, statements (1), (2) and (3) are all equivalent where,

(3) there is an one-to-one continuous function $g : \Omega \to (0, 1)$ such that, $g \in E$ and given any topological vector space V and a continuous function $f : \Omega' \to V$, f is uniformly approximable by g.

Proof. (1) Suppose equation (1) holds. Let α be an ordinal ≥ 1 such that,
$E = (f_\beta : \beta$ ordinal, $\beta < \alpha)$ where each $f_\beta \in C(\Omega', R)$ Let $\beta < \alpha$.
There exists $K_\beta \in (0, \infty)$ such that, $\sup_{x \in \Omega'} |f_\beta(x)| \leq K_\beta$.
For all $x \in \Omega'$, define $f'_\beta(x) = (8K_\beta)^{-1}(f_\beta(x) + K_\beta)$. Let
$E' = \{f'_\beta : \beta$ ordinal, $\beta < \alpha\}$. Let $x \in \Omega'$. Define

(1.1) $$f(x) = (f'_\beta(x) : \beta \text{ Ordinal}, \beta < \alpha)$$

Equip $(R)^\alpha$ with the topology $\tau_{\alpha, (R)^\alpha}$ (cf. equation (2.7) in section 2). $f : \Omega' \to (R)^\alpha$ is an one-to-one continuous function. By some known theorems (cf. [6], p.111 and 131),

(1.2) $f(\Omega')$ is a compact subset of $(R)^\alpha$ and for every open subset A of Ω', there exists $A' \in \tau_{\alpha,(R)^\alpha}$ such that,

$$f(A) = A' \cap f(\Omega')$$

Using equation (1.2) and theorem 2, one can easily show that,each $h \in C(\Omega', R)$ is (Ω, τ', E') -trigonometrically approximable and so, $E = C(\Omega', R)$.

(2) Suppose now (Ω', τ') is a Hausdorff, compact, second countable topological space and statement (2) holds. Since (Ω', τ') is a Hausdorff compact topological space, there exists a metric d on Ω' such that, d induces the topology τ' on Ω'. Let $(x_0, x_1, x_2, \cdots\cdots)$ be a countable dense subset of Ω. For each integer n ≥ 0 and for all $x \in \Omega'$, define $f_n(x) = d(x, x_n)$. Let $E' = (f_0, f_1, f_2, \cdots\cdots)$. For all $x \in \Omega$ define $h(x) = (f_0(x), f_1(x), f_2(x), \cdots)$. $h : \Omega' \to (R)^\omega$ is an one-to-one continuous function. Further, $h(\Omega)$ is a compact subset of $(R)^\omega$, and for every open subset A of Ω, there is an open subset A' of $(R)^\omega$ such that, $A = A' \cap h(\Omega)$. By theorem 2, statement 3 in the statement of theorem 3 holds. The fact that, statement 3 in the statement of theorem 3 implies statement 1 in the statement of theorem 3 is trivial.

The Stone-Weierstrass approximation theorem (cf.[6], p.160) states that, if E be a nonempty closed vector subspace of $C(\Omega', R)$ which contains constant functions, which is an algebra and which seperates points of Ω, $E = C(\Omega, R)$ Since for every real x, both Cos x and Sin x admits of power-series expansions, Stone-Weierstrass approximation theorem follows from theorem 3.

REFERENCES

1. CRAMER, HARALD. (1962) :"Mathematical methods of Statistics", Asia Publishing House, Bombay.

2. GHOSH, S. (1991) : "Some results related to characterisation of distribution functions and some applications", Proceedings of the 1991 international A.M.S.E. Conference on signals, data and systems held in New Delhi, India during December 9-11, 1991, Volume I, p. 19-31, A.M.S.E. Press, France.

3. HALMOS, P.R. (1964) : "Measure theory", Van Nostrand, New York.

4. NEVEU, J. (1965) : "Mathematical Foundations of the calculus of Probability, Holden Day, San-Fransisco.

5. RUDIN, W. (1974) : " Real and complex analysis", Tata McGraw Hill, New Delhi, India.

6. SIMMONS, G.F. (1963) : "Introduction to Topology and modern analysis", McGraw Hill, New York.

7. SZE-TSEN, TU. (1979) : " Introduction to General Topology", Tata McGraw Hill Publishing Company Limited, New Delhi.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 67-78

DEVELOPMENT OF 'RENAME DIRECTORY' COMMAND IN MS-DOS

K. J. Satao
Assistant Professor
Deptt. of Computer Sc. and Engg.,
B. N. College of Engg.,
PUSAD - 445215. Distt. - Yavatmal
Maharashtra(INDIA)

Abstract :

The operating system MS-DOS has the following commands for the directories- CHDIR(CD), MKDIR(MD), and RMDIR(RD) i.e. there is no command in MS-DOS (like RENDIR) which could rename a directory. There is a command called RENAME(REN) in MS-DOS, but this command renames the first file as a second file. Even though the directory name is stored in the directory area on the disk as a file name, this directory name cannot be renamed by the RENAME command. If we try to rename a directory name by the RENAME command we get an error message "Invalid path or file name". I am giving, herewith, a program which can act as a command to rename a directory and can be a good enhancement to the MS-DOS operating system.

1.0 Existing Procedure to Rename a Directory :

Directly renaming a directory name is not possible in MS-DOS. The MS-DOS manual gives the following procedure to rename a directory which is very much error prone. The procedure for renaming a directory as per the manual is as under...

1> Create a new directory.
2> Copy the files from the old directory to the new directory.
3> Delete the contents of the old directory.
4> Remove the old directory.

The problems with this procedure are as under...

1> If the root directory area on the disk becomes full then we cannot create a new directory on the disk.

2> Normally for copying the files from the old directory to the new directory we use the command copy *.* to the new directory. This process duplicates the files which may create a problem of the data area on the disk becoming full.

In this cases we have to stop copying and search the unwanted files which could be deleted.

But the biggest problem is copy *.* does not copy the hidden and system files.It can copy the read only files but their attributes become archive in the new directory.

3> If there is a subdirectory within a subdirectory then the copy command copies only the subdirectory name and not the files in it.

4> For deleting the contents of the old directory we normally use del *.* which creates a problem if there are hidden/system/read only files in the directory because DEL *.* cannot delete these files.

5> RMDIR does not work unless it is made empty.

In order to solve the above problems I have developed a program which will rename a directory without following the above procedure.The program is written in Turbo Pascal 5.0 under MS-DOS 4.0.

2.0 Introduction :

Before we go into the details of renaming a directory it is worthwhile to understand as to how the directories are stored on the disk, rather for that matter what is the logical structure of a disk. Here we consider 360 KB DSDD floppy diskettes.

Every MS-DOS diskette is divided into four separate areas. These are...
1> Boot area/Boot sector,
2> File Allocation Table(FAT),
3> Directory area/Root,
4> Data area/Data space.
Let us now look briefly into these four areas.

2.1 The Boot Area :

The Boot area i.e. the boot sector is the 0th sector on the disk and contains two things, namely, Boot parameters and the Disk bootstrap program.

The purpose of the disk bootstrap program is to load MS-DOS files residing on the diskette into the memory. The boot parameters are used by O.S. any time a read or write operation is performed on the disk. The boot parameters are...

Description	Length	Disk Bytes	Value
Jump instruction	3 bytes	EB 3C 90	EB 3C 90
System ID	8 bytes	4D 53 44 4F 53 34 2E 30	MSDOS4.0
No.of bytes per sector	2 bytes	00 02	512
No.of sectors per cluster	1 byte	02	2
No.of sectors in reserved area	2 byte	01 00	1
No.of copies of FAT	1 byte	02	2
No.of root directory entries	2 bytes	70 00	112
Total no. of sectors	2 bytes	D0 02	720
Media descriptor	1 byte	FD	FD
No.of sectors per FAT	2 bytes	02 00	2
No.of sectors per track	2 bytes	09 00	9
No.of sides	2 bytes	02 00	2
No.of hidden sectors	2 bytes	00 00	0

Of the above parameters, apart from the first 11 bytes rest is called as BIOS parameter block (BPB).

2.2 File Allocation Table :

The FAT occupies 4 sectors(sectors 1 to 4) on the disk after the boot area. The FAT maps the usage of all the disk space in the data space of the disk. It contains the information about the space used by each individual file, the unused disk space and the space that is unusable due to defects in the disk. Since FAT contains vital information, two copies of FAT are stored on the disk, just in case one gets destroyed, the other can be used.
A FAT entry can contain any of the following :

- Unused cluster
- Reserved cluster
- Bad cluster
- Last cluster in the file
- Next cluster number in the file

There is one entry in the FAT for each cluster in the data area. If the value in a FAT entry doesn't mark an unused, reserved or defective cluster, then the cluster corresponding to the FAT entry is a part of a file, and the value in the FAT entry would indicate the next cluster in the file.

This means that the space that belongs to a given file is mapped by a chain of FAT entries, each of which points to the next entry in the chain. The first cluster number in the chain is the starting cluster number in the file's directory entry. When a file is created or extended, MS-DOS allocates clusters(discussed later) to the file by searching the FAT for unused clusters and adding them to the chain. One cluster at a time is allocated to a file including a subdirectory entry. Vice versa, when a file is deleted, MS-DOS frees the clusters that had been allocated to the file or a subdirectory entry by clearing the corresponding FAT entries. The FAT chain for a file ends with an entry FFFF in the FAT.

2.3 The Root Directory :

The root directory usually occupies 7 sectors after FAT(sectors 5 to 11).The root directory contains a series of 32 byte entries. Each of these 32 byte entries are either a filename, a subdirectory name or the volume label of the disk. If it's a file entry the 32 byte entry contains information about file's size, it's starting location on the disk, it's size, and it's date and time of creation. This information is present in the following order.

Description	Size	Format	Disk Bytes	Value
1.Filename	8 bytes	ASCII	44 49 52 30 30 30 30 31	DIR00001
2.Extension	3 bytes	ASCII	44 41 54	DAT
3.Attribute	1 byte	bit coded	10	16
4.Reserved for future use	10 bytes	unused	ALL ZEROS	ALL BLANKS
5.Time	2 bytes	encoded	98 04	12:36a
6.Date	2 bytes	encoded	E3 1A	07-03-93
7.Starting cluster number	2 bytes	integer	60 01	352
8.Size	4 bytes	long INT	ALL ZEROS	ALL BLANKS

The attribute of a file can be...read only = 1 or hidden = 2 or system file= 4 or volume ID = 8 or directory = 16 or archive = 32 or a combination of these e.g.IO.SYS is a read only,hidden,system file. A subdirectory can also have such a attribute i.e. a subdirectory can be read only or hidden or system or archive or a combination of these. For a subdirectory entry also one cluster is allocated.

In the attribute byte each bit represents either the type of the file or whether the entry is a volume label entry or a sub directory name entry. The meaning of each bit is as shown below :

Bit numbers								Meaning
7	6	5	4	3	2	1	0	
							1	Read only
						1		Hidden
					1			System
				1				Volume label entry
			1					Sub directory
		1						Archive bit
	1							Unused
1								Unused

The starting cluster number indicates the place where the file begins on the disk data space. For volume label entry the starting cluster number is 0. For the directory entry which is a sub directory and contains another files, the root directory entry contains a starting cluster no. of another set of directory entries which reside in the disk data space e.g. If DIR1 is a subdirectory in the root directory which contains DIR2 as a subdirectory along with the other files, and DIR2 contains DIR3 as a subdirectory along with the other files, and DIR3 contains a set of files then the root directory entry will contain only one entry for DIR1 which will give the address of DIR1 on the disk data space where all it's files will be located; in DIR1 there will be one entry for DIR2 which will give the address of DIR2 on the disk data space; DIR2 will contain one entry for DIR3 which will give the address of DIR3 on the disk data space, and on this address all the files of DIR3 will be located.

2.4 The Data Space :

All files' data and subdirectories are stored in this area. It occupies the last and the largest part of each disk.

MS-DOS allocates space to files one cluster at a time. A cluster is nothing but a group of sectors that MS-DOS allocates to a file at a time or a subdirectory entry or a volume label.

As a file is being created, or an existing file is extended, the file's allocated space grows. When more space is needed, MS-DOS allocates another cluster to the file.

3.0 Program's Logic :

We can compile the program, given herewith, in the Turbo Pascal 5.0 environment and generate an executable file, called RENDIR.EXE. This program can be executed from the MS-DOS prompt in two ways. In the first way we can give just the name of the program viz. A>RENDIR <enter>. In the second way we can have A>RENDIR OLD_DIRECTORY NEW_DIRECTORY <enter>. If RENDIR.EXE exists in the root directory and we are in some other path then we can invoke the program by \RENDIR <enter> e.g. A:\K1\K2\K3>\RENDIR <enter>.

If we invoke the program without any parameters or parameters not equal to 2 then the program displays the following message and asks us to give the old directory name...

```
THIS PROGRAM IS FOR RENAMING A DIRECTORY...ENTER OLD DIRECTORY NAME DR:\PATH
\DIRNAME(MAX.65 CHARS)... OR ..\..\..\.. FORM  OR ..\..\..\..\DIRECTORY NAME
FORM
```

The user is supposed to give the old directory name. The program then asks..

```
ENTER NEW DIRECTORY NAME DR:\PATH\DIRNAME(MAX.65 CHARS)  OR ..\..\..\.. FORM
OR ..\..\..\..\DIRECTORY NAME FORM  DEFAULT PATH OF THE SOURCE DIRECTORY IS
ALLOWED USE THE SAME DRIVE...
```

The user is supposed to give the new directory name.

If the parameters are given then the parameters are assigned to the old directory name and the new directory name.

Both the directory names are first converted into the capital letters. The first routine executed is the PATH_CHECK. This routine checks whether both the previous paths(prior to the given directories) are same or not. We need not give the path name explicitly in both the directory names. We can give the path name only in one of the directory names, the program can take the other path name from this. This routine also converts both the directory names into the 8 + 3 format i.e. 8 characters primary name and 3 characters secondary name. If we want to rename the current directory then we can give just . for the old directory name e.g.

A:\K1\K2\K3>\RENDIR . KK <enter>. This will rename the current directory to KK and will give the proper prompt viz. A:\K1\K2\KK>. We can use ..\..\.. combination for both the directory names. But a proper care has to be taken while using this combination in the new directory name e.g. we can use

A:\K1\K2\K3>\RENDIR ..\.. ..\..\..\kk. This will rename K1 to KK and will give the prompt A:\KK\K2\K3>. If the prompt is A:\K1\K2\K3> and K3 contains K4 and we want to rename K4 then we can use

A:\K1\K2\K3>\RENDIR K4 K5. The renaming will be done and we will have the same prompt. The drive letter can be given or it can be omitted e.g. we can use A:\K1\K2\K3>\RENDIR A:K1 KK or A:\K1\K2\K3>\RENDIR \K1 KK or
A:\K1\K2\K3>\RENDIR A:K1 A:KK or A:\K1\K2\K3>\RENDIR \K1 A:KK or
A:\K1\K2\K3>\RENDIR \K1 \K2 etc.

We can rename any other directory out of this path also, but a proper path has to be given at least in one of the directory names.

If the paths of the old directory and the new directory don't match then the program gives the message...

DIFFERENT PATH/DRIVE IS NOT ALLOWED...COMMAND ABORTED...and the program ends.

If the path names are the same then the program executes the routine ADD_DRIVE_NAME which adds the drive name A: or B: or C: or D: or E: to the path of the directory name if it is not specified. After this the CHECK_CUR_DIR routine is executed. This routine is required because we cannot rename a directory when we are currently in the same directory. If we want to rename the current directory then we have to go to it's previous directory in the same path. This routine does this task.

After this the WILD_CHECK routine and the DIR_NAME_CHECK routines are executed for the old directory name and the new directory name. The WILD_CHECK routine checks whether the wild cards * and ? are present in the directory name. It also checks whether blank is present in the directory name. We can create a directory having a blank in it's name e.g. DIR 01, but it cannot be accessed by the MS-DOS CD or CHDIR command, so blank is not allowed in the directory name. If any of the directory names contain(in 8 + 3) these characters i.e. * or ? or blank then the program gives the message...

WILD CARDS ? & * AND BLANKS ARE NOT ALLOWED IN A DIRECTORY NAME...COMMAND ABORTED... and the program ends.

The DIR_NAME_CHECK routine checks that none of the directory names are null or the MS-DOS file reserved words viz. COM1 or COM2 or LPT etc.. If the directory name is null then the program gives the message...

DIRECTORY NAME CANNOT BE NULL...COMMAND ABORTED... and the program ends. If the directory name is one of the MS-DOS reserved file names then the program gives the message...

DIRECTORY NAME CANNOT BE... AUX/ CLOCK$/ COM1/ COM2/ COM3/ COM4/ CON/ LPT1/ LPT2/ LPT3/ LST/ NUL/ PRN...COMMAND ABORTED... and the program end.

After this the DIR_CHECK_1 routine is executed which checks whether the old directory name exists or not, if it exists then it is a directory or not.

This routine checks all the possible(16) attributes of a directory i.e. read only or hidden or system or archive directory or a combination of these. If the old directory name is not a directory or it does not exist then the program gives the message...

INNAME,' DOES NOT EXIST OR NOT A DIRECTORY...COMMAND ABORTED...'and the program ends.

After this the DIR_CHECK_2 routine is executed which checks whether the new directory name is a directory itself. If so then the program gives the message...

OUTNAME' IS A DIRECTORY ITSELF...COMMAND ABORTED...' and the program ends.

After this the program executes the routine SAME_NAME which checks whether the old directory name and the new directory name are different or not. If they are not different then the program gives the message...

RENDIR CANNOT RENAME SAME DIRECTORY...COMMAND ABORTED... and the program ends.

After this the program executes the routine FILE_CHECK. This routine checks whether the new directory name exists on the disk as a file. If it exists as a file name on the disk, in the particular path, then the program gives the following message and asks the following question...

'FOUND...',OUTNAME,'...AS A FILE ON THE DISK'
'ERASE...',OUTNAME,'...TO CREATE IT AS A DIRECTORY ? (Y/y) '
If the user responds with Y/y then the program deletes the file from the disk so that a directory could be created as that name, otherwise the program gives the message...
'...COMMAND ABORTED...' and the program ends.

If the old directory name is the current directory name then the CHECK_CUR_DIR routine sets the value of the variable OD_EQ_CD as TRUE. If OD_EQ_CD is FALSE then the routine WITHIN_PATH is executed. This routine checks whether the old directory name is included in the current directory name. If it is so then this routine sets the value of the variable INPATH to be TRUE. This is required to change the prompt correctly.

After this the program converts the old directory name and the new directory name into the ASCIIZ string, and invokes the MS-DOS function request 56H which does the renaming. If the operation is successful then the program gives the message...

'DIRECTORY... ',INNAME,'...RENAMED AS DIRECTORY... ',OUTNAME and the program ends after giving a correct prompt. If the operation is not successful then the program gives the message...

'DIRECTORY NOT RENAMED...CHECK ACCESS RIGHTS...' and the program ends.

4.0 Acknowledgements :

The author is highly thankful to Prof. B. M. Thakare, Principal, B. N. College of Engineering, Pusad and Prof. D. J. Das, H.O.D., Computer Science and Engineering, B. N. College of Engineering, Pusad for making the facilities available for this work.

The author is highly indebted to Mr. R. S. Verma and Mr. S. J. Doiphode,both programmers,B. N. College of Engineering,Pusad and Prof. R. B. Patil of B. N. College of Engg.,Pusad for testing the work and for useful suggestions.

5.0 Program for the command :

```
PROGRAM RENDIR;
{PROGRAM FOR RENAME DIRECTORY BY PROF. K. J. SATAO,B.N.COLLEGE OF ENGG.,
 PUSAD - 445 215,   DIST. : YAVATMAL,   MAHARASHTRA(INDIA)}
USES DOS;
CONST
  {R=READ ONLY,H=HIDDEN,S=SYSTEM,A=ARCHIVE,D=DIRECTORY}
  RHSAD = 55;RHSD  = 23;RHAD  = 51;RHD   = 19;RSAD  = 53;RSD   = 21;
  RAD   = 49;RD    = 17;HSAD  = 54;HSD   = 22;HAD   = 50;HD    = 18;
  SAD   = 52;SD    = 20;AD    = 48;D     = 16;
  CARRY_FLAG= 1;
  DIRSET  : ARRAY [1..5] OF STRING[3]=('AUX','CON','LST','NUL','PRN');
  DIRSET1 : ARRAY [1..7] OF STRING[4]=('COM1','COM2','COM3','COM4',
                                       'LPT1','LPT2','LPT3');

 VAR
  F1,F2            : FILE;
  ATTR             : WORD;
  INNAME,OUTNAME   : STRING[72];
  INAME,ONAME      : STRING[72];
  CH               : CHAR;
  REG              : REGISTERS;
  I                : INTEGER;
  P1,P2            : PATHSTR;
  D1,D2            : DIRSTR;
  N1,N2            : NAMESTR;
  E1,E2            : EXTSTR;
  OD_EQ_CD,INPATH  : BOOLEAN;
  S,S1,S2          : STRING;
  DR_NO            : INTEGER;
  K1               : STRING[2];
  FOUND            : BOOLEAN;
 PROCEDURE DRIVE_NO(DIRNAME : STRING);
 BEGIN
   S1 := '';     S2 := '';
   S1 := COPY(DIRNAME, 1, 1);
   S2 := COPY(DIRNAME, 2, 1);
   IF S2 = ':' THEN
   BEGIN
     IF ((S1 = 'a') OR (S1 = 'A')) THEN DR_NO := 1;
     IF ((S1 = 'b') OR (S1 = 'B')) THEN DR_NO := 2;
     IF ((S1 = 'c') OR (S1 = 'C')) THEN DR_NO := 3;
     IF ((S1 = 'd') OR (S1 = 'D')) THEN DR_NO := 4;
     IF ((S1 = 'e') OR (S1 = 'E')) THEN DR_NO := 5;
   END
   ELSE
   DR_NO := 0;
 END;
 PROCEDURE PATH_CHECK;
 LABEL 1,2,3,4,5,10;
 VAR
   L,N  :  INTEGER;
 BEGIN
   {ROUTINE TO CHECK WHETHER THE PATH NAMES ARE SAME OR NOT}
   DRIVE_NO(INNAME);
   L := LENGTH(INNAME);
   GETDIR(DR_NO,S);
   IF ((L = 1) AND (INNAME[1] = '.')) THEN
   BEGIN
     INNAME := S; GOTO 5;
   END;
```

```
    I  := 1;    K1 := '';    N  := 0;
    IF INNAME[1] = '.' THEN GOTO 1 ELSE GOTO 5;
1:  K1 := INNAME[I] + INNAME[I+1];
    IF K1 = '..' THEN N := N + 1 ELSE GOTO 2;
    I := I + 3;
    GOTO 1;
2:  IF N > 0 THEN
    BEGIN
      P1  := S;
      FOR I := 1 TO N DO
      BEGIN
        D1 := '';
        FSPLIT(P1, D1, N1, E1);
        P1 := D1;
        DELETE(P1,LENGTH(P1),1);
      END;
      DELETE(INNAME,1,N*2+N-1);
      INNAME := D1 + INNAME;
      DELETE(INNAME,LENGTH(D1),1);
    END;
5:  DRIVE_NO(OUTNAME);
    L := LENGTH(OUTNAME);
    GETDIR(DR_NO,S);
    IF ((L = 1) AND (OUTNAME[1] = '.')) THEN
    BEGIN
      OUTNAME := S;
      GOTO 10;
    END;
    I  := 1;    K1 := '';    N  := 0;
    IF OUTNAME[1] = '.' THEN GOTO 3 ELSE GOTO 10;
3:  K1 := OUTNAME[I] + OUTNAME[I+1];
    IF K1 = '..' THEN N := N + 1 ELSE GOTO 4;
    I := I + 3;
    GOTO 3;
4:  IF N > 0 THEN
    BEGIN
      P1  := S;
      FOR I := 1 TO N DO
      BEGIN
        D1 := '';
        FSPLIT(P1, D1, N1, E1);
        P1 := D1;
        DELETE(P1,LENGTH(P1),1);
      END;
      DELETE(OUTNAME,1,N*2+N-1);
      OUTNAME := D1 + OUTNAME;
      DELETE(OUTNAME,LENGTH(D1),1);
    END;
10: IF INNAME[1]  = '\' THEN INNAME  := S[1] + S[2] + INNAME;
    IF OUTNAME[1] = '\' THEN OUTNAME := S[1] + S[2] + OUTNAME;
    P1 := INNAME;
    FSPLIT(P1, D1, N1, E1);
    INNAME := D1 + N1 + E1;
    P2 := OUTNAME;
    FSPLIT(P2, D2, N2, E2);
    OUTNAME := D2 + N2 + E2;
    IF (D1 = '') THEN
    BEGIN
      D1 := D2;
      INNAME := D1 + INNAME;
    END;
```

```
  IF (D2 = '') THEN
  BEGIN
    D2 := D1;
    OUTNAME := D1 + OUTNAME;
  END;
  IF D1 <> D2 THEN
  BEGIN
    WRITELN('DIFFERENT PATH/DRIVE IS NOT ALLOWED...COMMAND ABORTED...');
    CHDIR(S);HALT;
  END;
END;
PROCEDURE ADD_DRIVE_NAME;
BEGIN
  S2 := '';     S2 := COPY(D1, 2, 1);
  IF S2 <> ':' THEN
  BEGIN
    DRIVE_NO(INNAME);
    GETDIR(DR_NO,S);
    S1 := '';    S1 := COPY(S, 1, 2);
    INNAME  := S1 + INNAME;
  END;
  S2 := '';
  S2 := COPY(D2, 2, 1);
  IF S2 <> ':' THEN
  BEGIN
    DRIVE_NO(OUTNAME);
    GETDIR(DR_NO,S2);
    S1 := '';    S1 := COPY(S2, 1, 2);
    OUTNAME := S1 + OUTNAME;
  END;
END;
PROCEDURE CHECK_CUR_DIR;
BEGIN
  {ROUTINE TO CHECK WHETHER THE SOURCE DIR IS THE CURRENT DIR}
  OD_EQ_CD := FALSE;
  DRIVE_NO(INNAME);
  S := '';
  GETDIR(DR_NO,S);
  P1 := ''; D1 := ''; N1 := ''; E1 := '';
  IF S = INNAME THEN
  BEGIN
    OD_EQ_CD := TRUE;
    P1       := INNAME;
    FSPLIT(P1, D1, N1, E1);
    DELETE(D1,LENGTH(D1),1);
    CHDIR(D1);
  END;
END;
PROCEDURE WILD_CHECK(DIRNAME : STRING);
BEGIN
  {ROUTINE TO CHECK WILD CARDS * & ? AND BLANKS}
  FOR I:= 1 TO LENGTH(DIRNAME) DO
  BEGIN
    IF ((DIRNAME[I]='?')OR(DIRNAME[I]='*')OR(DIRNAME[I]=' ')) THEN
    BEGIN
      WRITE('WILD CARDS ? & * AND BLANKS ARE NOT ALLOWED IN A ');
      WRITE('DIRECTORY NAME...COMMAND ABORTED...');
      WRITELN;CHDIR(S);HALT;
    END;
  END;
END;
```

```
PROCEDURE DIR_NAME_CHECK(DIRNAME : STRING);
BEGIN
  FOUND := FALSE;
  IF (DIRNAME = '') THEN
  BEGIN
    WRITELN('DIRECTORY NAME CANNOT BE NULL...COMMAND ABORTED...');
    CHDIR(S);HALT;
  END;
  {ROUTINE FOR INVALID DIRECTORY NAMES}
  FOR I := 1 TO 5 DO
  BEGIN
    IF DIRNAME = DIRSET[I] THEN FOUND := TRUE;
  END;
  FOR I := 1 TO 7 DO
  BEGIN
    IF DIRNAME = DIRSET1[I] THEN FOUND := TRUE;
  END;
  IF DIRNAME = 'CLOCK$' THEN FOUND := TRUE;
  IF FOUND THEN
  BEGIN
    WRITE('DIRECTORY NAME CANNOT BE...');
    WRITE('AUX/CLOCK$/COM1/COM2/COM3/COM4/CON/LPT1/LPT2/LPT3/LST/');
    WRITE('NUL/PRN...COMMAND ABORTED...'); WRITELN; CHDIR(S); HALT;
  END;
END;
PROCEDURE DIR_CHECK_1;
BEGIN
  {ROUTINE TO CHECK THE SOURCE NAME  EXISTS OR NOT}
  {AND WHETHER IT IS A DIRECTORY OR NOT}
  ASSIGN(F1,INNAME);  GETFATTR(F1, ATTR);
  IF ((ATTR<>RHSAD) AND (ATTR<>RHSD) AND (ATTR<>RHAD) AND (ATTR<>RHD) AND
      (ATTR<>RSAD)  AND (ATTR<>RSD)  AND (ATTR<>RAD)  AND (ATTR<>RD)  AND
      (ATTR<>HSAD)  AND (ATTR<>HSD)  AND (ATTR<>HAD)  AND (ATTR<>HD)  AND
      (ATTR<>SAD)   AND (ATTR<>SD)   AND (ATTR<>AD)   AND (ATTR<>D)) THEN
  BEGIN
    WRITE(INNAME,' DOES NOT EXIST OR NOT A DIRECTORY ');
    WRITELN('...COMMAND ABORTED...');CHDIR(S);HALT;
  END;
END;
PROCEDURE DIR_CHECK_2;
BEGIN
  {ROUTINE TO CHECK WHETHER DESTINATION NAME EXISTS}
  ASSIGN(F2,OUTNAME); GETFATTR(F2, ATTR);
  IF ((ATTR=RHSAD) OR (ATTR=RHSD) OR (ATTR=RHAD) OR (ATTR=RHD) OR
      (ATTR=RSAD)  OR (ATTR=RSD)  OR (ATTR=RAD)  OR (ATTR=RD)  OR
      (ATTR=HSAD)  OR (ATTR=HSD)  OR (ATTR=HAD)  OR (ATTR=HD)  OR
      (ATTR=SAD)   OR (ATTR=SD)   OR (ATTR=AD)   OR (ATTR=D)) THEN
  BEGIN
    WRITELN(OUTNAME,' IS A DIRECTORY ITSELF...COMMAND ABORTED...');
    CHDIR(S);HALT;
  END;
END;
PROCEDURE SAME_NAME;
BEGIN
  {ROUTINE TO CHECK SAME DIRECTORY NAMES}
  IF INNAME = OUTNAME THEN
  BEGIN
    WRITELN('RENDIR CANNOT RENAME SAME DIRECTORY...COMMAND ABORTED...');
    CHDIR(S);HALT;
  END;
END;
```

```
PROCEDURE FILE_CHECK;
BEGIN
  {ROUTINE TO CHECK WHETHER DESTINATION FILE EXISTS}
  {$I-}
  RESET(F2);
  {$I+}
  IF IORESULT = 0 THEN
  BEGIN
    WRITELN('FOUND...', OUTNAME,'...AS A FILE ON THE DISK');
    CLOSE(F2);
    WRITE('ERASE...',OUTNAME, '...TO CREATE IT AS A DIRECTORY ? (Y/y) ');
    READLN(CH);
    IF UPCASE(CH) = 'Y' THEN
       ERASE(F2)
    ELSE
    BEGIN
      WRITELN('...COMMAND ABORTED...');
      CHDIR(S);HALT;
    END;
  END;
END;
PROCEDURE WITHINPATH;
BEGIN
  {ROUTINE TO CHECK WHETHER THE SOURCE DIR IS A SUBSET OF CUR DIR}
  INPATH := FALSE;  S := '';
  DRIVE_NO(INNAME);
  GETDIR(DR_NO,S);
  S1 := ''; S1 := COPY(S, 1, LENGTH(INNAME));
  IF S1 = INNAME THEN INPATH := TRUE;
END;
BEGIN
  {MAIN ROUTINE}
  INPATH   := FALSE; OD_EQ_CD := FALSE; FOUND    := FALSE;
  IF PARAMCOUNT <> 2 THEN
  BEGIN
    WRITE('THIS PROGRAM IS FOR RENAMING A DIRECTORY...');
    WRITE('ENTER OLD DIRECTORY NAME DR:\PATH\DIRNAME(MAX.65 CHARS)...');
    WRITE('OR ..\..\..\.. FORM OR ..\..\..\..\DIRECTORY NAME FORM');
    WRITELN;
    READLN(INNAME);
    FOR I := 1 TO LENGTH(INNAME) DO INNAME[I] := UPCASE(INNAME[I]);
    WRITE('ENTER NEW DIRECTORY NAME DR:\PATH\DIRNAME(MAX.65 CHARS)');
    WRITE('OR ..\..\..\.. FORM OR ..\..\..\..\DIRECTORY NAME FORM');
    WRITE('( DEFAULT PATH OF THE SOURCE DIRECTORY IS ALLOWED)');
    WRITE(' USE THE SAME DRIVE...');
    WRITELN;
    READLN(OUTNAME);
    FOR I := 1 TO LENGTH(OUTNAME) DO OUTNAME[I] := UPCASE(OUTNAME[I]);
    PATH_CHECK;
    ADD_DRIVE_NAME;
    CHECK_CUR_DIR;
    WILD_CHECK(INNAME);
    DIR_NAME_CHECK(INNAME);
    WILD_CHECK(OUTNAME);
    DIR_NAME_CHECK(OUTNAME);
    DIR_CHECK_1;
    DIR_CHECK_2;
    SAME_NAME;
    FILE_CHECK;
    IF OD_EQ_CD = FALSE THEN WITHINPATH;
  END
  ELSE
  BEGIN
    INNAME := PARAMSTR(1);
    FOR I := 1 TO LENGTH(INNAME)  DO INNAME[I]  := UPCASE(INNAME[I]);
    OUTNAME := PARAMSTR(2);
    FOR I := 1 TO LENGTH(OUTNAME) DO OUTNAME[I] := UPCASE(OUTNAME[I]);
    PATH_CHECK;
    ADD_DRIVE_NAME;
    CHECK_CUR_DIR;
    WILD_CHECK(INNAME);
    DIR_NAME_CHECK(INNAME);
    WILD_CHECK(OUTNAME);
    DIR_NAME_CHECK(OUTNAME);
    DIR_CHECK_1;
    DIR_CHECK_2;
    SAME_NAME;
    FILE_CHECK;
    IF OD_EQ_CD = FALSE THEN WITHINPATH;
  END;
  INAME   := INNAME;
  ONAME   := OUTNAME;
  INNAME  := INNAME  + CHR(0);
  OUTNAME := OUTNAME + CHR(0);
  WITH REG DO
  BEGIN
    DS := SEG(INNAME[1]);
    DX := OFS(INNAME[1]);
    AH := ($56);
    ES := SEG(OUTNAME[1]);
    DI := OFS(OUTNAME[1]);
    MSDOS(REG);
    IF ((FLAGS AND CARRY_FLAG) =1 ) THEN
    BEGIN
       WRITELN('DIRECTORY NOT RENAMED...CHECK ACCESS RIGHTS...');
       CHDIR(S);HALT;
    END
    ELSE
    BEGIN
      WRITELN('DIRECTORY...   ',INNAME,
                 '...RENAMED AS DIRECTORY... ',OUTNAME);
      IF OD_EQ_CD THEN CHDIR(OUTNAME);
      IF INPATH THEN
      BEGIN
        GETDIR(DR_NO,S); DELETE(S,1,LENGTH(INAME));
        S := ONAME + S;  CHDIR(S);
      END;
    END;
  END;
END.
```

6.0 References :

1> Turbo Pascal 5.0 User's Guide.
2> Turbo Pascal 5.0 Reference Guide.
3> MS-DOS User's Guide and User's Reference.
4> Programmer's Reference manual for IBM Personal Computer
By Steven Armbrust & Ted Forgeron.
5> The MS-DOS Handbook
By Richard Allan King.
6> LET US C
By Yashwant Kanetkar.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 79-90

DEVELOPMENT OF 'SEARCH' AND 'REPLACE' COMMANDS IN MS-DOS USING A LINKED LIST METHOD

K. J. Satao
Assistant Professor
Deptt. of Computer Sc. and Engg.,
B. N. College of Engineering,
PUSAD - 445215. Distt. - Yavatmal
Maharashtra(INDIA)

Abstract :

The operating system MS-DOS contains a command called FIND(FIND.EXE) for searching a particular string in a file or a group of files. This is not a full fledged command. It is a filter(a command that reads our input, transforms in some way, then outputs to our screen) and has some limitations. Turbo Pascal 5.0 has given a command GREP.EXE(similar to UNIX's grep command). But neither MS-DOS nor UNIX has a command to replace a string throughout a file system(directory) or a group of files. I am giving , in this paper, a program which can act as a command to search and replace a string throughout a directory or a group of files which can be a good enhancement to the MS-DOS operating system.

1.0 Introduction

The FIND command of MS-DOS looks for a string in one or more files. After searching, the command displays any lines it has found that contain the specified string.

The limitations of this command are...

1> The uppercase characters in the string will not match the lowercase characters
2> If the string contains quotation marks, we must enclose it in additional pair of quotes.
3> Wild cards (* and ?) are not allowed in filenames or extensions.

I have developed a command using a linked list method which removes the above limitations and has many other advantages. The GREP command supplied by Turbo Pascal, though very much versatile, is available to only those users who have purchased Turbo Pascal 5.0, whereas my command, I, hereby authorize any user of computers to use.

There are commands in MS-DOS viz. REPLACE.EXE and SUBST.EXE. The purpose of this replace command is to update the previous version of the files and that of the subst command is to substitute a string for a path. In nutshell these commands(replace and subst) do not allow us to replace /substitute a given string by another string in files i.e. there is no command in MS-DOS (like find - which searches a given string throughout the file system) to replace a given string by another string throughout the file system.

My program can replace a given string throughout the file system using a linked list method. My program takes care of both the equal and the unequal strings. The string up to 20 characters can be searched and can be replaced by another string of up to 20 characters.

The command can be a good enhancement to MS-DOS operating system because no such command exists in MS-DOS. The program can be converted into the UNIX environment to enrich UNIX utilities as well.

The program is written in Turbo Pascal 5.0 under MS-DOS 4.01.

2.0 Advantages of the program

1> The uppercase characters in a string will match the lowercase characters and vice-versa, if ignore case is opted for.
2> If a string contains quotation marks, there is no need to enclose it in another pair of quotes.
3> ? and * are allowed as wild cards in the string, if opted for.
4> ? and * are allowed as wild cards in the file names and extensions.
5> We can replace by query, if opted for.
6> A string up to 20 characters can be searched throughout the file system i.e. the directory.
7> The string can be searched/replaced in any path/directory on the disk.
8> The program works under Novell's NETWARE operating i.e. in a local area network as well.
9> Any normal(archive) file e.g. .EXE,.BAK,.DOC,.C,.PAS,.COB,.FOR,.PRG,.DBF, .WK1, W.S. Document file,W.S. Non document file,.COM etc. can be searched and replaced.
10 The program takes care of both equal and unequal strings and the multiple occurrences of the string in a file or a group of files.
11 A string can be replaced by null i.e. removed from a file/files.
12 We can abort search/replace while in operation, if opted for.
13 We can search/replace even the non alphabetic characters.

3.0 Program's Logic

I have used a linked list method of data structures. My program has facilities like S for Search, R for Replace, and Q for Quit.

The command for running my program is -

```
     DR-NAME>srandrep filename<enter>  OR  DR-NAME>srandrep<enter>.
```

After the command is executed we get a screen as below...

```
-----------------------------------------------------------------------
WELCOME TO THE SEARCH AND REPLACE ROUTINES  BY PROF.K.J.SATAO, C.O.E.,PUSAD,
INDIA
-----------------------------------------------------------------------
YOUR COMMAND PL. S-SEARCH/R-REPLACE/Q-QUIT
```

If the command is not r/R/s/S/q/Q then the program once again asks the command. If the command is q/Q then the program displays the message...

'SEE YOU AGAIN...BYE...' and the program halts giving the MS-DOS prompt.

If the file name is not given at the MS-DOS prompt then the program asks...
'PLEASE GIVE THE FILE NAME(79 CHARS MAX)e.g.C:\TP\KJS\PAPER*.* '.

If we want to search in the current directory then the full path need not be given, just the file name e.g. test*.* is sufficient.
This file name is converted into the uppercase and is tested for it's validity. If the file name is given at the DOS prompt then it is converted into the uppercase and the validity of the file name is tested.

For testing the file name validity the following procedure is used ...
If the file name is null or the file name is one of the following 13 names AUX, CLOCK$, COM1, COM2, COM3, COM4, CON, LPT1, LPT2, LPT3, LST, NUL, PRN then the program gives the message... 'INVALID FILE NAME...'and the program asks the valid file name.

If the file name is a valid file name then the program checks the existence of at least one occurrence of the file using the FINDFIRST call.

The program takes care of the following types of errors...
'FILE NOT FOUND...', 'PATH NOT FOUND...', 'ACCESS DENIED...',
'INVALID HANDLE...','NOT ENOUGH MEMORY...','INVALID ENVIRONMENT...',
'INVALID FORMAT...','NO MORE FILES...'.

In case any of the above errors occur then the program aborts the operation and once again asks the command.

In case there is at least one occurrence for the file then the program proceeds further and gets the options for ignore case and wild cards by asking the following queries...
'DO YOU WANT TO IGNORE CASE ? y/Y FOR YES...'
'? OR * CAN BE USED AS A WILD CHARACTER...USE IT? y/Y FOR YES...'.

After this the program asks the string to be searched or replaced by asking the question...
'WHICH STRING?(MAX 20 CHARACTERS) '.

If the ignore case is opted for then this string is converted into capital letters. We cannot search or replace a null string.

If the command is for searching then the following steps are followed...

1> Find the first occurrence of the file using the FINDFIRST call.
2> Read the file using the routines READ_NEXT_FILE and READLINES. The READ_NEXT_FILE routine gets the actual filename along with path,prepares for new links and calls the READLINES routine. In this routine we read either up to linefeed or up to 128 characters, each time from the file and insert into a linked list of 166 characters,till the size of the file is reached. All the characters from the file,including CR,LF,EOF etc. are inserted into the linked list. If a particular file is too big (no. of lines>2870 or the size of the file > 367360 bytes) then the program gives the message...
'NOT ENOUGH MEMORY TO LOAD THE FILE...USE OTHER METHOD TO SEARCH/REPLACE'
and the processing for the other files continues.

At the end we get a message...
'TOTAL LINES = 'total lines' TOTAL CHARACTERS = 'total characters in the file.
3> Give the message...'SEARCHING IS GOING ON...'
4> Execute SEARCH routine. This routine scans the linked list for the particular string. If a particular line contains the string the program finds the no. of occurrences in the line and gives a message...
' NO. OF OCCURRENCES IN THIS LINE = 'actual occurrences. If the string is found in many lines then after each such 5 lines the program gives the message...
'STRIKE ANY KEY TO CONTINUE...OR s/S TO STOP...' and depending upon the option it proceeds further.
At the end the program gives the message...

' TOTAL OCCURRENCES = ' total occurrences
'STRIKE ANY KEY TO CONTINUE...OR a/A TO ABORT...'.
If the option is not 'a/A' then the program goes to step 5 otherwise it aborts the searching operation and goes to step 7.
5> Find the next file to be searched using the FINDNEXT routine.
6> The steps 2 to 5 are repeated till all the files are exhausted.
7> At the end the program gives the message...
'TOTAL NO. OF FILES SEARCHED = ' actual number
'TOTAL NO. OF FILES HAVING THE STRING = 'actual number
'STRIKE ANY KEY TO CONTINUE...'
and the control goes to the command options.

If the command is for replacing then the following steps are followed...

1> The program gets the string, by which to be replaced, by asking the following question and getting a replay for it...
'BY WHICH STRING?(MAX 20 CHARACTERS) '.
2> The program then asks the following question and gets a response for it..
'REPLACE BY QUERY ? (y/Y) '.
3> Same as step 1 of searching.
4> Same as step 2 of searching.
5> Give the message...'REPLACEMENT IS GOING ON...'
6> Execute REPLACE routine. This routine scans the linked list for the particular string. If a particular line contains the string then,if query is opted for then the program asks for the response...
'DO YOU WANT TO REPLACE? y/Y FOR YES...a/A TO ABORT...'
and acts according the choice. If the string is found and query is not opted for,then direct replacement is done. The program takes care of the multiple occurrences of the string in any line and equal as well as unequal strings.

All the replacement are first done in the linked list. After the process is over then the program displays the message...
'TOTAL OCCURRENCES = 'total occurrences in the file.
The program then writes the contents of the linked list into the same file in the same path and gives the following message and gets a response for it...
'REPLACEMENTS OVER...STRIKE ANY KEY TO CONTINUE...OR a/A TO ABORT'.

If the option is not 'a/A' then the program goes to step 7 otherwise it aborts the replacement operation and goes to step 9.
7> Find the next file to be replaced using the FINDNEXT routine.
8> The steps 4 to 7 are repeated till all the files are exhausted.
9> At the end the program gives the message...
'TOTAL NO. OF FILES SEARCHED = ' actual number
'TOTAL NO. OF FILES REPLACED = ' actual number
'STRIKE ANY KEY TO CONTINUE...'
and the control goes to the command options.

4.0 Limitations of the work

My program has the following limitations...

1> We cannot search/replace CR,LF,EOF, AND NULL characters.
2> Files having no. of lines > 2870 or the size of the file > 367360 bytes cannot be searched or replaced if the main memory is 640 KB.
3> For a line of 128 characters we can have...
    a> 128 occurrences for 1 character getting replaced by 1 character.
    b> 38 occurrences for 1 character getting replaced by 2 characters.
    c> 19 occurrences for 1 character getting replaced by 3 characters.
    and so on
    d> 2 occurrences for 1 character getting replaced by 20 characters.

5.0 Further enhancements under consideration

I plan to undertake the following improvements in my present program...

1> Search/Replace the entire disk by finding the tree structure.
2> Search/Replace the system/hidden/readonly file. This is very easy to incorporate. We have to just change the attribute of the file before the processing and after the processing.
3> Search/Replace without using the linked list method.Of course this method will also have some limitations.

6.0 The Program

```
PROGRAM SEARCH_AND_REPLACE;
{PROGRAM FOR SEARCH AND REPLACE BY PROF. K.J. SATAO, B.N.COLLEGE OF ENGG.,'}
{PUSAD - 445 215,   DIST. : YAVATMAL,     MAHARASHTRA(INDIA)}
USES CRT,DOS;
LABEL 1,2,3,4,5;
CONST
   FILESET  : ARRAY [1..13] OF STRING[6]=('AUX','CLOCK$','COM1','COM2',
              'COM3','COM4','CON','LPT1','LPT2','LPT3','LST','NUL','PRN');
TYPE
   TEXT1    = ARRAY[1..166] OF CHAR;
   STRING12 = ARRAY[1..20]  OF CHAR;
   POINTER1 = ^ENTRY;
   ENTRY    =
   RECORD
      LINENO        : INTEGER;
      LINETEXT      : TEXT1;
      NEXT          : POINTER1;
   END;
VAR
   CASE1,QUERRY,WILD,COMMAND,CH  :  CHAR;
   START,STOP,HIGHNO:INTEGER;NEWLINE:TEXT1;
   HEAD,LINK,PREV,P: POINTER1;TEMP,TEMP1,TEMP2,B   :   TEXT1;  F  : TEXT;
   TEMP12,TEMP13,TEMP14,B12,B13:STRING12;BUF : ARRAY[1..10240] OF CHAR;
   M,N,O,I,J,K,I1,J1,K1,NU,OCC,TOCC,L,LINEOCC,COUNT2,COUNT1 :  INTEGER;
   DONE : BOOLEAN; FILENAME,FILNAME : STRING[79]; P1 : POINTER;
   DIRINFO : SEARCHREC; TCH : LONGINT;
   D : DIRSTR; N1 : NAMESTR; E : EXTSTR;PARCOUNT : WORD;
PROCEDURE GETNAME;
BEGIN
   WRITE('PLEASE GIVE THE FILE NAME(79 CHARS MAX)');
   WRITE('e.g.C:\TP\PAPER12.PAS  '); READLN(FILENAME);
   WRITELN;
END;
PROCEDURE FILETEST;
LABEL 15;
BEGIN
   DONE := FALSE;
   IF (FILENAME = '') THEN
   BEGIN
   WRITELN('INVALID FILE NAME...');
   DONE := TRUE;
   GOTO 15;
   END;
   FOR I:=1 TO LENGTH(FILENAME) DO
   BEGIN
   CH := FILENAME[I];
   IF ((ORD(CH)>=97) AND (ORD(CH)<=122)) THEN
   FILENAME[I]:=CHR(ORD(FILENAME[I])-32);
   END;
   FOR I:=1 TO 13 DO
   BEGIN
   IF FILENAME = FILESET[I] THEN
   BEGIN
   WRITELN('INVALID FILE NAME...');
   DONE := TRUE;
   GOTO 15;
   END;
   END;
 15:END;
```

```
PROCEDURE CHECK_ERROR;
BEGIN
   IF DOSERROR = 2 THEN WRITELN('FILE NOT FOUND...') ELSE
   IF DOSERROR = 3 THEN WRITELN('PATH NOT FOUND...') ELSE
   IF DOSERROR = 5 THEN WRITELN('ACCESS DENIED...')  ELSE
   IF DOSERROR = 6 THEN WRITELN('INVALID HANDLE...') ELSE
   IF DOSERROR = 8 THEN WRITELN('NOT ENOUGH MEMORY...')   ELSE
   IF DOSERROR =10 THEN WRITELN('INVALID ENVIRONMENT...') ELSE
   IF DOSERROR =11 THEN WRITELN('INVALID FORMAT...')      ELSE
   IF DOSERROR =18 THEN WRITELN('NO MORE FILES...');
END;
PROCEDURE WRITELINES;
LABEL 71,72;
BEGIN
   ASSIGN(F,FILNAME);ERASE(F);SETTEXTBUF(F,BUF);REWRITE(F);LINK:=HEAD^.NEXT;
   WHILE (LINK <> NIL)  DO
   BEGIN
   TEMP := LINK^.LINETEXT;
   FOR I := 166 DOWNTO 1 DO
   BEGIN
   IF TEMP[I]=CHR(0) THEN GOTO 71;GOTO 72;
71:END;
72:FOR NU := 1 TO I DO
   BEGIN
   WRITE(F,TEMP[NU]);
   END;
   LINK := LINK^.NEXT;
   END {WHILE};
   WRITELN; CLOSE(F); WRITELN;
   WRITE('REPLACEMENT OVER...STRIKE ANY KEY TO CONTINUE...OR a/A TO ABORT');
   CH := UPCASE(READKEY);
END {WRITELINES};
PROCEDURE INSERT(NUMBER: INTEGER; VAR NEWLINE: TEXT1);
BEGIN
   DONE := FALSE; PREV := HEAD; LINK := HEAD^.NEXT;
   WHILE (LINK <> NIL) AND NOT DONE DO
   BEGIN
   IF LINK^.LINENO = NUMBER THEN
   BEGIN
   NEW(P); P^.LINENO   := NUMBER;
   P^.LINETEXT := NEWLINE; P^.NEXT := LINK; PREV^.NEXT  := P;
   REPEAT
   NU := LINK^.LINENO;  NU := NU+1; LINK^.LINENO:= NU;
   LINK := LINK^.NEXT;
   UNTIL LINK  = NIL; DONE  := TRUE;
   END {IF}  ELSE
   IF LINK^.LINENO > NUMBER THEN
   BEGIN
   NEW(P);P^.LINENO := NUMBER; P^.LINETEXT := NEWLINE; P^.NEXT := LINK;
   PREV^.NEXT  := P; DONE   := TRUE;
   END {IF}  ELSE
   BEGIN
   PREV  := LINK; LINK  := LINK^.NEXT;
   END;
   END {WHILE};
   IF NOT DONE THEN
   BEGIN
   NEW (P); P^.LINENO  := NUMBER; P^.LINETEXT:= NEWLINE;
   P^.NEXT    := NIL; PREV^.NEXT :=P;
   END
END {INSERT};
```

```
PROCEDURE READLINES(SNUMBER : INTEGER);
LABEL 10;
BEGIN
   ASSIGN(F, FILNAME); SETTEXTBUF(F, BUF);RESET(F);
   NU := 0;TCH := 0;
   FOR I := 1 TO 166 DO TEMP1[I] := CHR(0);
   WHILE (NOT (TCH = DIRINFO.SIZE)) DO
   BEGIN
   READ(F,CH);NU:=NU+1;TEMP[NU]:=CH;
   TCH := TCH + 1;
   IF ((TEMP[NU] = CHR(10)) OR (NU > 127)) THEN
   BEGIN
   TEMP1[NU] := CH;
   IF MEMAVAIL < 182 THEN
   BEGIN
   WRITE('NOT ENOUGH MEMORY TO LOAD THE FILE...');
   WRITELN('USE OTHER METHOD TO SEARCH/REPLACE...');
   WRITE('STRIKE ANY KEY TO CONTINUE...');
   CH := READKEY; GOTO 10;
   END;
   NEWLINE := TEMP1;INSERT(SNUMBER,NEWLINE);
   SNUMBER := SNUMBER + 1;
   FOR K := 1 TO 166 DO TEMP1[K] := CHR(0);
   HIGHNO:=HIGHNO+1;NU := 0;
   END
   ELSE TEMP1[NU] := TEMP[NU];
   END;
   IF TEMP[NU-1]<>CHR(10) THEN
   BEGIN
   NEWLINE := TEMP1;INSERT(SNUMBER,NEWLINE);
   SNUMBER := SNUMBER + 1;
   HIGHNO:=HIGHNO+1;
   END;
10:CLOSE(F);
END;
PROCEDURE LIST(START, STOP: INTEGER);
LABEL 30,31;
BEGIN
   DONE  := FALSE; LINK  := HEAD^.NEXT;
   WHILE (LINK <> NIL) AND NOT DONE DO
   BEGIN
   IF LINK^.LINENO > STOP THEN DONE  := TRUE
   ELSE
   IF LINK^.LINENO >= START THEN
   BEGIN
   WRITE(LINK^.LINENO:4,':');
   TEMP  := LINK^.LINETEXT;
   FOR  I := 166 DOWNTO 1 DO
   BEGIN
   IF TEMP[I]=CHR(0) THEN GOTO 30; GOTO 31;
30:END;
31:FOR I1 := 1 TO I-1 DO
   BEGIN
   CH := TEMP[I1]; WRITE(CH);
   END;
   CH := TEMP[I]; WRITELN(CH);
   LINK := LINK^.NEXT
   END {IF}
   ELSE LINK := LINK^.NEXT;
   END {WHILE};
END;
```

```
PROCEDURE CHECK_IG_CASE_WILD;
BEGIN
   FOR J := 0 TO L-1 DO
   BEGIN
   TEMP14[J+1] := TEMP[I1+J];
   IF CASE1 = 'Y' THEN
   BEGIN
   IF ((ORD(TEMP14[J+1])>=97) AND (ORD(TEMP14[J+1])<=122)) THEN
       TEMP14[J+1] := CHR(ORD(TEMP14[J+1])-32);
   END;
   END;
   IF WILD = 'Y' THEN
   BEGIN
   FOR J:= 1 TO L DO
   IF TEMP12[J] = '?' THEN TEMP14[J] := '?';
   IF TEMP12[J] = '*' THEN TEMP14[J] := '*';
   END;
END;
PROCEDURE READ_NEXT_FILE;
BEGIN
   MARK(P1);NEW(HEAD);HEAD^.NEXT := NIL;HIGHNO :=1;
   FILNAME := D + DIRINFO.NAME;
   WRITELN;TEXTCOLOR(YELLOW+BLINK);WRITELN(DIRINFO.NAME);TEXTCOLOR(YELLOW);
   WRITELN;READLINES(HIGHNO); COUNT1 := COUNT1 + 1;
   WRITELN('TOTAL LINES = ',HIGHNO-1,
           '             TOTAL CHARACTERS = ',DIRINFO.SIZE:7);
   START := 1; STOP  := HIGHNO-1;
END;
PROCEDURE SEARCH(START, STOP: INTEGER;VAR TEMP12 : STRING12);
LABEL 30,31,32,33;
BEGIN
   OCC:=0;TOCC:=0;K1:=0;DONE:=FALSE;LINK := HEAD^.NEXT;
   WHILE (LINK <> NIL) AND NOT DONE DO
   BEGIN
   IF LINK^.LINENO > STOP THEN DONE := TRUE
   ELSE
   IF LINK^.LINENO >= START THEN
   BEGIN
   TEMP := LINK^.LINETEXT;
   FOR I:= 166 DOWNTO 1 DO
   BEGIN
   IF TEMP[I]=CHR(0) THEN GOTO 30; GOTO 31;
30:END;
31:I1  := 1;
   WHILE I1<=I DO
   BEGIN
   CHECK_IG_CASE_WILD;
   IF TEMP14 <> TEMP12 THEN
   BEGIN
   I1 := I1 +1; GOTO 32;
   END;
   OCC  := OCC+1; TOCC := TOCC+1; I1   := I1 + L;
32:END;
   IF OCC>0 THEN
   BEGIN
   K1   := K1 +1; WRITELN(LINK^.LINENO:5,':',TEMP);
   WRITELN('     NO. OF OCCURANCES IN THIS LINE = ',OCC:5);
   K    :=  K1 DIV 5;
   IF  K1   = K*5 THEN
   BEGIN
   WRITELN;
   WRITE('STRIKE ANY KEY TO CONTINUE...OR s/S TO STOP...');
   CH := UPCASE(READKEY);
   WRITELN;
   IF CH = 'S' THEN GOTO 33;
   END;
   END;
   END;{IF}
   LINK := LINK^.NEXT; OCC  := 0;
   END {WHILE};
33:WRITELN;
   WRITELN('     TOTAL OCCURANCES = ',TOCC:5);
   IF TOCC > 0 THEN
   BEGIN
   WRITE('STRIKE ANY KEY TO CONTINUE...OR a/A TO ABORT...');
   CH := UPCASE(READKEY); COUNT2 := COUNT2 + 1;
   END;
END;
PROCEDURE REPLACE(START, STOP: INTEGER;VAR TEMP12,TEMP13 : STRING12);
LABEL 30,31,32,50,51,52;
BEGIN
   OCC:= 0; TOCC:= 0; DONE:= FALSE; LINK:= HEAD^.NEXT;
   WHILE (LINK <> NIL) AND NOT DONE DO
   BEGIN
   IF LINK^.LINENO > STOP THEN DONE  := TRUE
   ELSE
   IF LINK^.LINENO >= START THEN
   BEGIN
   TEMP  := LINK^.LINETEXT; TEMP2 := LINK^.LINETEXT;LINEOCC:=0;
   FOR I := 166 DOWNTO 1 DO
   BEGIN
   IF TEMP[I]=CHR(0) THEN GOTO 30;
   GOTO 31;
30:END;
31:I1  := 1;
51:WHILE I1<=I DO
   BEGIN
   CHECK_IG_CASE_WILD;
   IF TEMP14 <> TEMP12 THEN
   BEGIN
   I1 := I1 +1; GOTO 32;
   END;
   OCC  := OCC+1;LINEOCC := LINEOCC + 1; TOCC := TOCC+1;
   IF QUERRY = 'Y' THEN
   BEGIN
   WRITELN(LINK^.LINENO:5,':',TEMP2);
   WRITE('DO YOU WANT TO REPLACE? y/Y FOR YES...a/A TO ABORT...');
   CH := UPCASE(READKEY);WRITELN;
   IF CH = 'Y' THEN GOTO 50;
   IF CH = 'A' THEN GOTO 52;
   I1 := I1 + L; GOTO 32;
   END;
50:IF L = M THEN
   BEGIN
   FOR J := 0 TO L-1 DO TEMP2[I1+J] := TEMP13[J+1];
   I1  := I1 + L;
   END;
   IF L<M THEN
   BEGIN
   J1 := I1+(OCC-1)*(M-L);
   FOR J := 0 TO L-1 DO TEMP2[J1+J] := TEMP13[J+1]; J  := J1+L;
   FOR K := I+(M-L)*(LINEOCC-1) DOWNTO J  DO TEMP2[K+M-L] := TEMP2[K];
   J  := J1;
   FOR K :=  L+1 TO M DO
   BEGIN
   TEMP2[J+L]:= TEMP13[K]; J  := J+1;
   END;
   I1  := I1 + L;
   END;
   IF L>M THEN
   BEGIN
   J1 := I1-(OCC-1)*(L-M);
   FOR J := 0 TO M-1 DO TEMP2[J1+J] := TEMP13[J+1];
   J := (I1 + L) - (L-M)*OCC;
   FOR K := I1+L TO I  DO
   BEGIN
   TEMP2[J] := TEMP[K]; J := J+1;
   END;
   FOR K := I DOWNTO I-LINEOCC*(L-M)+1 DO TEMP2[K] := CHR(0);
   I1  := I1 + L;
   END;
32:END;
   IF OCC>0 THEN LINK^.LINETEXT := TEMP2;
   OCC  := 0; LINK := LINK^.NEXT;LINEOCC := 0;
   END {IF}
   ELSE LINK := LINK^.NEXT;
   END {WHILE};
   LINK := LINK^.NEXT;
   WRITELN; WRITELN('TOTAL OCCURANCES = ',TOCC:5);
   IF TOCC> 0 THEN
   BEGIN
   COUNT2 := COUNT2 + 1; WRITELINES;
   END;
52:END;
BEGIN
   PARCOUNT := PARAMCOUNT;
1: CLRSCR;TEXTCOLOR(YELLOW);
   WRITE('--------------------------------------------------------------');
   WRITE('--------------------'); TEXTCOLOR(RED);
   WRITE('WELCOME TO THE SEARCH AND REPLACE ROUTINES  BY PROF.K.J.SATAO,C');
   WRITE('.O.E.,PUSAD,INDIA'); TEXTCOLOR(YELLOW);
   WRITE('--------------------------------------------------------------');
   WRITE('--------------------'); TEXTCOLOR(YELLOW);
   WRITE('YOUR COMMAND PL.');
   TEXTCOLOR(CYAN);WRITE('S');TEXTCOLOR(YELLOW);WRITE('-SEARCH/');
   TEXTCOLOR(CYAN);WRITE('R');TEXTCOLOR(YELLOW);WRITE('-REPLACE/');
   TEXTCOLOR(CYAN+BLINK);WRITE('Q');TEXTCOLOR(YELLOW);WRITE('-QUIT ');
   READLN(COMMAND);WRITELN;TEXTCOLOR(YELLOW);WRITELN;
   COMMAND := UPCASE(COMMAND);
   IF ((COMMAND<>'S') AND (COMMAND<>'R') AND (COMMAND<>'Q')) THEN GOTO 1;
   IF COMMAND = 'Q' THEN
   BEGIN
   WRITELN; TEXTCOLOR(CYAN+BLINK);WRITELN('SEE YOU AGAIN...BYE...');
   WRITELN; TEXTCOLOR(YELLOW);GOTO 2;
   END;
   {ROUTINES SEARCH AND REPLACE START}
   IF PARCOUNT = 0 THEN  GETNAME ELSE FILENAME := PARAMSTR(1);
   FILETEST;PARCOUNT :=0;
   IF DONE THEN
   BEGIN
   WRITELN;WRITE('INVALID FILE NAME...STRIKE ANY KEY TO CONTINUE...');
   CH := READKEY;GOTO 1;
   END;
   FINDFIRST(FILENAME,ARCHIVE,DIRINFO);
   CHECK_ERROR;  FSplit(FILENAME, D, N1, E);
   IF DOSERROR<>0 THEN
   BEGIN
   WRITELN('STRIKE ANY KEY TO CONTINUE...'); CH := READKEY; GOTO 1;
   END;
   COUNT2 := 0; COUNT1 := 0;
   WRITE('DO YOU WANT TO IGNORE CASE ? y/Y FOR YES...');
   READLN(CASE1); CASE1 := UPCASE(CASE1);
   WRITELN;WRITE('? OR * CAN BE USED AS A WILD CHARACTER...');
   WRITE('USE IT? y/Y FOR YES...');
   READLN(WILD); WILD := UPCASE(WILD);
   WRITELN;
3: WRITELN;WRITE('WHICH STRING?(MAX 20 CHARACTERS) ');
   FOR I := 1 TO 20 DO
   BEGIN
   TEMP12[I] := CHR(0); TEMP14[I] := CHR(0);
   END;
   L := 0;
   WHILE NOT EOLN DO
   BEGIN
   L := L+1; READ(TEMP12[L]);
   END;READLN;
   IF L =0 THEN GOTO 3;
   IF CASE1 = 'Y' THEN
   BEGIN
   FOR I:=1 TO L  DO
   BEGIN
   CH := TEMP12[I];
   IF ((ORD(CH)>=97) AND (ORD(CH)<=122)) THEN
       TEMP12[I] := CHR(ORD(ORD(CH)-32));
   END;
   END;
   B12 :=TEMP12;WRITELN;
   IF COMMAND = 'S' THEN
   BEGIN
   FINDFIRST(FILENAME,ARCHIVE,DIRINFO);
   WHILE DOSERROR = 0 DO
   BEGIN
   CH := 'N'; READ_NEXT_FILE;
   WRITELN('SEARCHING IS GOING ON...');
   SEARCH(START,STOP,B12);
   RELEASE(P1);
   IF CH = 'A' THEN GOTO 4;
   FINDNEXT(DIRINFO);
   END;
4: WRITELN;WRITELN('TOTAL NO. OF FILES SEARCHED = ',COUNT1);WRITELN;
   WRITELN('TOTAL NO. OF FILES HAVING THE STRING = ',COUNT2);WRITELN;
   WRITE('STRIKE ANY KEY TO CONTINUE...');CH := READKEY; GOTO 1;
   END {CASE FOR S};
   IF COMMAND = 'R' THEN
   BEGIN
   WRITE('BY WHICH STRING?(MAX 20 CHARACTERS) ');
   FOR I := 1 TO 20 DO TEMP12[I] := CHR(0); M   := 0;
   WHILE NOT EOLN DO
   BEGIN
   M := M+1; READ(TEMP12[M]);
   END;
   READLN; B13 := TEMP12; WRITELN;
   WRITE('REPLACE BY QUERRY ? (y/Y) ');
   READLN(QUERRY);QUERRY := UPCASE(QUERRY);
   FINDFIRST(FILENAME,ARCHIVE,DIRINFO);
   WHILE DOSERROR = 0 DO
   BEGIN
   CH := 'N'; READ_NEXT_FILE;
   WRITELN('REPLACEMENT IS GOING ON...');
   REPLACE(START,STOP,B12,B13); RELEASE(P1);
   IF CH = 'A' THEN GOTO 5;
   FINDNEXT(DIRINFO);
   END;
5: WRITELN;WRITELN('TOTAL NO. OF FILES SEARCHED = ',COUNT1);WRITELN;
   WRITELN('TOTAL NO. OF FILES REPLACED = ',COUNT2);WRITELN;
   WRITE('STRIKE ANY KEY TO CONTINUE...');CH := READKEY; GOTO 1;
   END {CASE FOR R};
2: END.
```

7.0 References

1> Introduction to Data Structures with Pascal
By Bhagat Singh & Thomas L. Naps.
2> Masturing Turbo Pascal
By Tom Swan.
3> Turbo Pascal 5.0 User's Guide.
4> Turbo Pascal 5.0 Reference Guide.
5> An Interactive Line Text Editor for MS-DOS Operating System
-A paper by the author(Prof. K. J. Satao)presented at the Indian Computing Congress - 1993 at Hyderabad(INDIA) organized by the Chartered Computer Professionals of India.

8.0 Acknowledgements :

The author is highly thankful to Prof. B. M. Thakare, Principal, B. N. College of Engineering, Pusad and Prof. D. J. Das, H.O.D., Computer Science and Engineering, B. N. College of Engineering, Pusad for making the facilities available for this work.

The author is highly indebted to Mr. R. S. Verma and Mr. S. J. Doiphode,both programmers,B. N. College of Engineering,Pusad for testing the work and for useful suggestions.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 91-100

A PHONEME IDENTIFICATION SYSTEM

TESSAMMA THOMAS C.S. SRIDHAR
Department Of Electronics
Cochin University Of Science and Techonology
Cochin - 682 022, India.

ABSTRACT: A knowledge based phoneme identification method is presented in this paper. All the 36 phonemes present in the British English language are considered. The salient features of the speech signals, like short time energy (STE), short time zero-crossing rate (STZCR), correlation coefficients, pitch period, etc. have been made use of, for the identification. The fact that voiced phonemes like vowels and nasals have higher energy than other phonemes, forms the basis for the classification. Based on the average STE content in the phoneme, the 36 phonemes are first categorised into three groups. Within each group, a range of values are fixed up for the various parameters, and a final identification is done.

INTRODUCTION: A phoneme is the basic unit of any word in any language, which describes how speech conveys a linguistic meaning. Roughly speaking, a phoneme is a group of similar, but not identical, sounds that differ from one another in accordance with the context in which each occurs [1].

The air we inhale in, is pushed out of the lungs through the trachea, past the vocal chords and into the pharynx, and finally comes out through the mouth or nose or both, and is perceived as speech. Based on the mode of vibration of the source of excitation, speech sounds are broadly classified as voiced and unvoiced. Depending on the position of the different articulators in the process of the production of the various sounds, they are subdivided to form ten groups (as shown in fig.1), containing altogether 36 phonemes [2].

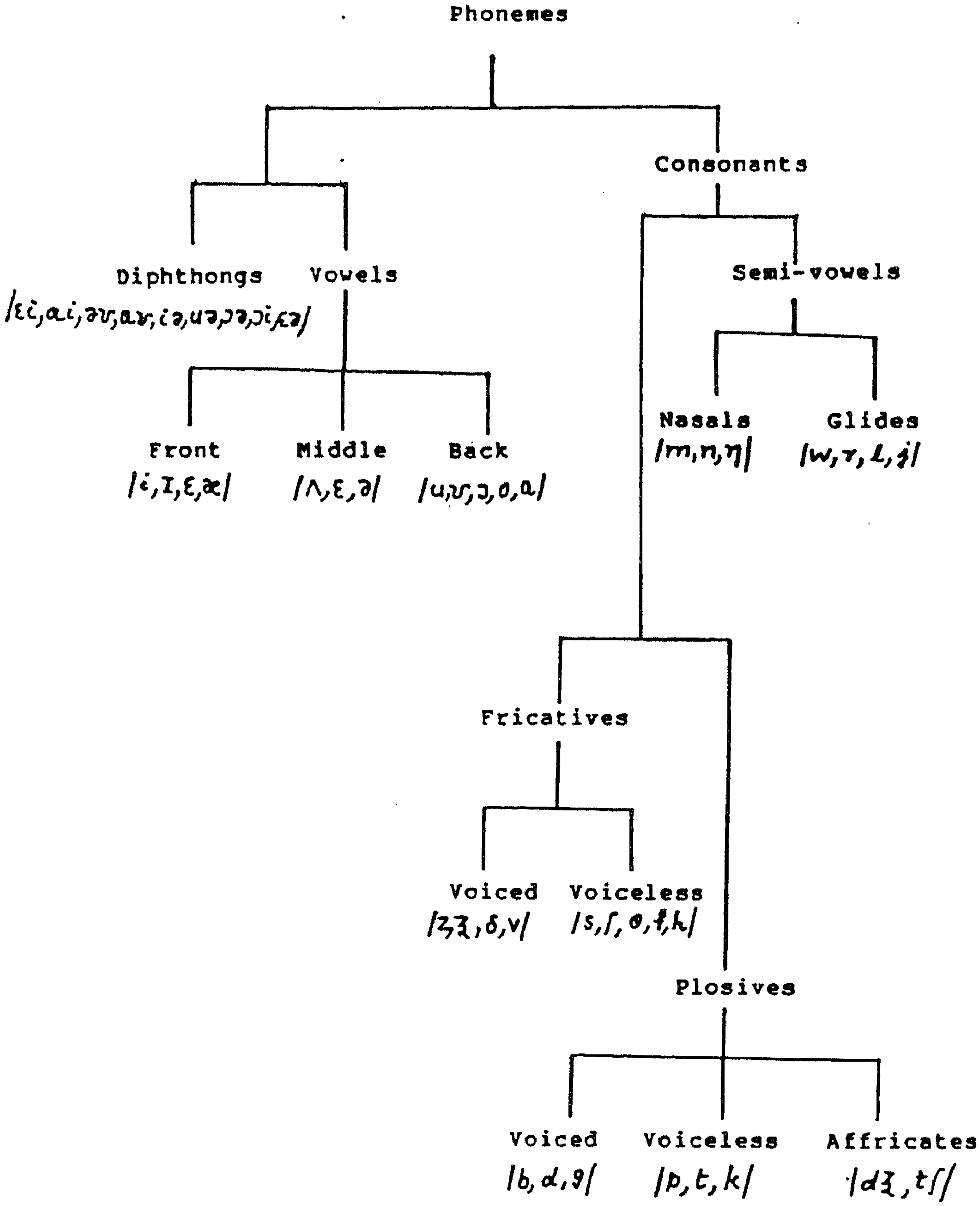


Fig.1 The Phonemes of British English [2]

FEATURE EXTRACTION FOR CREATING A KNOWLEDGE BASE

Three to four words each, were chosen for each of the 36 phonemes. The region corresponding to the particular phoneme was manually separated from each word, by plotting on a VDU, and then re-checking it by reconstruction and listening. The data corresponding to each phoneme were stored in different files.

It is reported that [4,5,6], the vocal tract 'rings' for a duration of about 10 ms. For a sampling frequency of 8 kHz, it corresponds to 80 samples. Hence, the speech data were grouped into segments of length 80 samples. The parameters like, STE, STZCR, STZCR-energy product (ZEP), STZCR to STE ratio (ZER), correlation coefficients for lags 1 and 2, ρ_1 and ρ_2, etc. were evaluated for each segment and a set of average values were calculated for each selected phoneme, in the different words chosen. From these, a range of values were obtained for each parameter, for each of the 36 phonemes. A list of the values of the different parameters useful in the phoneme identification process, is given in Table 1.

Studying the energy ranges of the different phonemes, three different ranges were fixed, and the phonemes that fall in a particular range were grouped together. If a particular phoneme has energy values in any two different ranges, they are included in both the groups, and are finally identified based on the parameters chosen in that particular group.

The energy limits obtained in the study were 0.09 to 0.47 and the three different ranges chosen were 0.09 to 0.17, 0.17 to 0.23, and 0.23 to 0.47. The phonemes that come under the topmost energy group (Group-I), are mostly the nasals and the vowels. Group-II contains the higher energy plosives, glides and fricatives, and low energy vowels. The third energy group contains phonemes having comparitively lower correlation than the higher energy phonemes of Group I and II. Within each group, a final identification is done, by fixing up different ranges for the values of ZCR, ZEP, ZER, pitch-period, ρ_1 and the ratio

Table 1 Analysis results of all the 36 phonemes in British English

Phoneme (1)	STE (2)	STZCR (3)	ZEP (4)	ZER (5)	PP (6)	ρ_1 (7)	ρ_1/ρ_2 (8)
/b/	.13-.15	1140-1458	170-187	7670-11,390	4.875	.83-.865	1.3-1.9
	.19-.21	792-1150	151-180	4047-4146	4.625-5.00	.91-.94	1.1-1.3
/d/	.14-.16	1479-1536	210-240	7800-10300	5.0	.75-.80	1.0-1.013
/g/	.15-.157	1020-1220	145-190	6800-7750	4.875	.89-.90	1.02
/p/	.14-.163	1075-1350	150-220	7675-8310	--	.82-.85	1.65-1.90
/t/	.14	1321	180	9645	--	.87	1.16
	.18	979	174	5500	--	.905	1.40
/k/	.17-.183	658-792	116-145	4050-4350		.92-945	1.14-1.15
/w/	.13-.166	946-1016	127-170	6120-7035	4.3-4.625	.91-.915	1.2-1.3
/r/	.11-.125	670-830	71-104	6280-6650	4.75-4.875	.92-.94	1.15-1.2
/l/	.14-.15	1140-1625	160-250	7979-10,690	4.75-4.875	.79-.88	1.1-1.25
/j/	.196	875	171.5	4465	4.75	.913	1.127
/h/	.09-.13	1300-1500	122-201	11194-13873	--	.63-.93	1.19
/f/	.19-.20	645-750	125-150	3329-3740	--	.86-.89	1.10-1.2
/θ/	.194	1083	210-210	5574	--	.84	1.127
/s/	.23-.29	560-710	160-163	1960-3080	--	.91-.93	1.14-1.16
/ʃ/	.17-.24	571-1812	94-193	3420-3460	--	.90-.93	1.09-1.13
/i/	.28-.29	530-535	150-190	1870-2320	4.125	.87	1.055
/ɪ/	.2-.227	525-646	119-131.14	2312-3185	4.125	.89-895	1.08-1.12
/ɛ/	.18-.22	833-875	155-189	4050-4450	4.625	.91	1.22-1.25
/æ/	.097-.14	1500-2375	204-230	10,980-24,540	5.0	.37-.60	1.3-2.14
/ʒ/	.18-.22	840-857	157-186	3958-4670	4.375-4.75	.93	1.2-1.21
	.24-.25	911-917	224-230	3615-3750	4.125-4.5	.93-.935	1.21
/ʌ/	.11-.115	1438-1650	157-190	13,192-14410	4.375-4.875	.67-.69	--
/ə/	.31	708	220.72	2270	4.125	.94	1.112
/a/	.095-.13	1060-1150	132-150	8496-8850	4.625-4.75	.635-.80	1.25-1.5
/o/	.162	857	138.62	5297	4.625	.923	1.238
/ɔ/	.25	1521	380.63	6076.7	4.5	.838	1.06
/ʊ/	.25-.45	479	119-213	1076-1925	4.0	.948-.978	1.06-1.13
/u/	.23-.274	429-458	105-117	1991-2045	4.125-4.375	.958-.97	1.08-1.12

/m/	.26-.46	480-489	189-221	1050-1275	4.875-5.0	.94-.97	1.05-1.09
/n/	.232-.42	446-460	187-195	1060-1955	4.375-4.75	.89-.99	1.05-1.3
/ŋ/	.34-.47	450-464	150-220	987-1340	4.125	.97	1.05-1.06
/ʋ/	.12-.13	859 -1281	100-172	7367-9540	4.875-5.25	.81-.93	1.2-1.33
/δ/	.205-.207	608-930	125-190	2930-4540	4.375-4.875	.94-.95	1.13-1.15
/z/	.14	804	113	5734	4.625	.84	1.377
	.199	518	103	2603	4.625-4.75	.95	1.48
	.234-258	429-562	110-129	1663-2444	4.5-4.625	.955-.965	1.16-1.26
/ʒ/	.200	953	190.6	4765	4.375-4.5	.925	1.082
/dʒ/	.24-.34	518-641	156-177	1514-2630	4.125-4.875	.93-.952	1.05-1.07
	.185	686	127	3700	4.75	.935	1.05
/tʃ/	.24-.325	500	120-163	1540-2075	4.125	.94-.96	1.05-1.09

Table 2

KNOWLEDGE-BASE FOR THE PHONEME IDENTIFICATION METHOD

(STE, STZCR (SHORT TIME ENERGY AND SHORT TIME ZEROCROSSING RATE OF THE INPUT SIGNAL))

(ZEP, ZER (ZEROCROSSING RATE ENERGY PRODUCT, AND ZEROCROSSING RATE BY ENERGY))

(K_1 (FIRST REFLECTION COEFFICIENT))

(PP (PITCH PERIOD OF THE INPUT SIGNAL))

(ρ_1, ρ_2 (NORMALISED CORRELATION COEFFICIENTS FOR LAGS 1 AND 2)

(ρ' (RATIO OF ρ_1 to ρ_2)

1 IF .23 ≤ STE ≤ .47
THEN GROUP I (ʒ, ɔ, ʃ, ə, ʂ, tʃ, m, n, ŋ, ʋ, u, dʒ, z, i)

2 IF GROUP I = YES AND STZCR = 1515 ± 10 OR ZER = 6000 ± 100
OR ZEP = 370 ± 10
THEN PHONEME /ɔ/

3 IF GROUP I = YES AND STZCR = 910 ± 100
OR ZEP = 225 ± 5
THEN PHONEME /ʒ/

4 IF GROUP I = YES AND STZCR = 810 ± 10

OR ZER = 3400 ± 100 OR ZEP = 195 ± 10
THEN PHONEME /ʃ/

5 IF GROUP I = YES AND STZCR = 710 ± 10
OR ZER = 2200 ± 100 OR ZEP = 215 ± 5
THEN PHONEME /ə/

6 IF GROUP I = YES AND STZCR = 630 ± 70
AND ρ_2 = .81 ± .01
THEN PHONEME /s/

7 IF GROUP I = YES AND STZCR = 580 ± 70
AND ρ_2 = .84 ± .01
THEN PHONEME /dʒ/

8 IF GROUP I = YES AND STZCR = 530 ± 5
AND ρ_1 = .87 ± .01
THEN PHONEME /i/

9 IF GROUP I = YES AND STZCR = 500 ± 5
AND ρ_1 = 0.93 ± .01
THEN PHONEME /tʃ/

10 IF GROUP I = YES and STZCR = 480 ± 5
AND PP = 4.875 ± .125 AND ZEP = 200 ± 20
THEN PHONEME /m/

11 IF GROUP I = YES and STZCR = 480 ± 2
AND PP = 4.0 ± .125 AND ZEP = 165 ± 50
THEN PHONEME /r/

12 IF GROUP I = YES and STZCR = 450 ± 10
AND PP = 4.25 ± .125
AND K_1 = -.96 ± .05
THEN PHONEME /u/

13 IF GROUP I = YES and STZCR = 490 ± 70
AND PP = 4.5 ± .125 AND K_1 = -.945 ± .05
THEN PHONEME /z/

14 IF GROUP I = YES AND STZCR = 455 ± 10
AND PP = 4.5 ± .25
THEN PHONEME /ŋ/

15 IF GROUP I = YES AND STZCR = 455 ± 10
AND PP = 4.125 ± 0
THEN PHONEME /ŋ/

16 IF .17 ≤ STE < .23
THEN GROUP II (t, θ, ʓ, j, ʒ, δ, k, f, ɪ, ɛ, z, dʓ)

17 IF GROUP II = YES AND STZCR = 1310 ± 20
OR ZER = 9650 ± 10 AND ρ_1 = .905 ± .05
THEN PHONEME /t/

18 IF GROUP II = YES AND STZCR = 1070 ± 20
OR ZER = 5550 ± 100 AND ρ_1 = .84 ± .05
THEN PHONEME /θ/

19 IF GROUP II = YES AND STZCR = 950 ± 20
OR ZER = 4750 ± 50 AND ρ_1 = .925 ± .05
THEN PHONEME /ʓ/

20 IF GROUP II = YES AND STZCR = 875 ± 20
AND ZER = 4460 ± 50 AND ρ_1 = .915 ± .05
THEN PHONEME /j/

21 IF GROUP II = YES AND STZCR = 850 ± 20
AND ZER = 4300 ± 350 AND ρ_1 = .930 ± .05
THEN PHONEME /ʒ/

22 IF GROUP II = YES AND STZCR = 815 ± 25
AND ZER = 4100 ± 50 AND ρ_1 = .925 ± .15
THEN PHONEME /b/

23 IF GROUP II = YES AND STZCR = 770 ± 160
AND ZER = 3750 ± 750 AND ρ_1 = .945 ± .05
THEN PHONEME /δ/

24 IF GROUP II = YES AND ZCR = 740 ± 60
AND ZER = 4150 ± 100 AND ρ_1 = .93 ± .15
THEN PHONEME /k/

25 IF GROUP II = YES AND STZCR = 700 ± 60
AND ZER = 3540 ± 200 AND ρ_1 = .875 ± .15
THEN PHONEME / f /

26 IF GROUP II = YES AND STZCR = 680 ± 20
AND ZER = 3700 ± 50 AND ρ_1 = .935 ± .05
THEN PHONEME / dʒ /

27 IF GROUP II = YES AND STZCR = 585 ± 60
AND ZER = 2750 ± 450 AND ρ_1 = .89 ± .05
THEN PHONEME /l/

28 IF GROUP II = YES AND STZCR = 530 ± 20
AND ZER = 1870 ± 20 AND ρ_1 = .87 ± .05
THEN PHONEME / ε /

29 IF GROUP II = YES AND STZCR = 510 ± 10
AND ZER = 2600 ± 20 AND ρ_1 = .925 ± .10
THEN PHONEME /z /

30 IF $.09 \leq$ STE $<$ .17
THEN GROUP III (t, w, r, v, z, o, b, d, g, p, l, h, æ , ʌ , a)

31 IF GROUP III = YES AND STZCR = 1940 ± 540 AND ZEP = 220 ± 15
OR ZER > 18000 AND ρ' = 1.85 ± .45
THEN PHONEME /æ /

32 IF GROUP III = YES AND STZCR = 1550 ± 100 AND ZEP = 175 ± 15
OR ZER = 13800 ± 600, AND ρ' = 2.75 ± .25
THEN PHONEME / ʌ /

33 IF GROUP III = YES AND STZCR = 1510 ± 30 AND ZEP = 225 ± 15
AND ZER = 10050 ± 2500 AND ρ' = 1.01 ± .01
THEN PHONEME / d /

34 IF GROUP III = YES AND STZCR = 1400 ± 100 AND ZEP = 160 ± 40
AND ZER = 12500 ± 1400 AND ρ' = 1.8 ± .1
THEN PHONEME /h/

35 IF GROUP III = YES AND STZCR = 1400 ± 250 and ZEP = 205 ± 45
AND ZER = 9300 ± 1300 AND ρ' = 1.15 ± .075
THEN PHONEME /ℓ/

36 IF GROUP III = YES AND STZCR = 1260 ± 200 AND ZEP = 140 ± 10
AND ZER = 11500 ± 3500 AND ρ' = 1.45 ± .02
THEN PHONEME /a/

37 IF GROUP III = YES AND STZCR = 1300 ± 160 AND ZEP = 180 ± 10
AND ZER = 9500 ± 2000 AND ρ' = 1.6 ± .3
THEN PHONEME /b/

38 IF GROUP III = YES AND STZCR = 1200 ± 150 AND ZEP = 185 ± 35
AND ZER = 8000 ± 330 AND ρ' = 1.8 ± .15
THEN PHONEME /p/

39 IF GROUP III = YES AND STZCR = 1120 ± 100 AND ZEP = 170 ± 25
AND ZER = 7300 ± 500, AND ρ' = 1.02 ± .01
THEN PHONEME /g/

40 IF GROUP III = YES AND STZCR = 1080 ± 220 AND ZEP = 135 ± 35
AND ZER = 8500 ± 1050 AND ρ' = 1.95 ± .25
THEN PHONEME /v/

41 IF GROUP III = YES AND STZCR = 980 ± 35 AND ZEP = 130 ± 44
AND ZER = 6570 ± 450 AND ρ' = 1.45 ± .15
THEN PHONEME /w/

42 IF GROUP III = YES AND STZCR = 980 ± 10 AND ZEP = 175 ± 5
AND ZER = 5500 ± 100 AND ρ' = 1.16 ± .05
THEN PHONEME /t/

43 IF GROUP III = YES AND STZCR = 855 ± 10 AND ZEP = 140 ± 5
AND ZER = 5300 ± 100 AND ρ' = 1.2 ± .05
THEN PHONEME /o/

44 IF GROUP III = YES AND STZCR = 805 ± 10 AND ZEP = 115 ± 5
AND ZER = 5750 ± 100 AND ρ' = 1.18 ± .05
THEN PHONEME /z/

45 IF GROUP III = YES AND STZCR = 750 ± 80 AND ZEP = 90 ± 20
AND ZER = 6450 ± 200 AND $\rho' = 1.175 \pm .025$
THEN PHONEME /γ/

$\rho' = \rho_1/\rho_2$. The detailed algorithm developed is shown in Table 2.

SIMULATION STUDY OF THE PHONEME IDENTIFICATION SYSTEM

Speech signals were low-pass filtered to 4 kHz, sampled at 8 kHz and encoded to 12 bits to form the data samples. The region corresponding to each phoneme was separated and stored in different files. Each set of data was then normalised, such that the maximum value of the samples in the set is unity. They were then grouped into blocks of 80 samples and a set of average values were obtained for the different parameters, in the case of each phoneme. The above algorithm was then applied. The recognition score was high, when data from the training sequences were used.

If more and more words are considered for each phoneme, the limiting ranges of the different parameters might change. Then the number of groups into which the phonemes were initially divided, based on STE, can be increased, and the same procedure be applied effectively. Further work has to be done to make the system speaker independent.

REFERENCES

[1] Frank Fallside and William A. Woods, Eds., Computer Speech Processing, Prentice Hall Inc., New Jersey, 1985.
[2] W.A.Ainsworth, Mechanisms of Speech Recognition, Pergamon Press Ltd., New York, 1976.
[3] L.R. Rabiner and R.W. Schafer, Digital Processing of Speech Signals, Prentice-Hall Inc., New Jersey, 1978.
[4] Tessamma Thomas and C.S.Sridhar,"A new method of adaptive linear prediction for voiced speech signals.", J.Acoust.Soc. India, vol.XVIII, (3 and 4), p.p.160-163,1990.
[5] M.R. Schoerder, "Linear predictive coding of speech.", IEEE, Communications Magazine, vol.23, no.8, 1985.
[6] B.S. Atal and S.L. Hanauer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave ", J. Acoust. Soc. Amer., Vol. 50, No. 2, pp. 637-655, 1971.
[7] Tessamma Thomas and Chetlur Sridhar, " A Knowledge-Based Speaker Recognition System.", AMSE periodical, Advances in Modelling and Analysis, B, vol.27, No.2, pp. 53-63, 1993.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 101-110

ADAPTIVE AUTOREGRESSIVE MODELLING FOR RESTORATION OF MUSIC

K.S.Jog, Satish S. Bhat, Murtaza Y. Nagarwala.
Electronics and Telecommunication Department
Government College of Engineering Pune, INDIA

ABSTRACT

In this paper determination of proper order for autoregressive model for a piece of music is described. The modelling is done in view of estimation of a piece of music for restoration. The signal is transformed in frequency domain. The modelling is then done for each spectral component i.e. bin. The order of the model for each bin is determined in an adaptive manner. Processing is done digitally through FFT and IFFT. The hardware used is commonly available PC-AT with sufficient RAM. The technique can be implemented on any other machine as well. Results give comparison of fixed order model, partially adaptive model and fully adaptive model.

1. INTRODUCTION

The signal being modelled is a piece of recorded music. Model is to be used for restoration of the signal in a small portion of recording which is either damaged or is deteriorated through heavy noise. Such portions are termed as breakages. The signal on either side of the breakage is assumed to be noisefree.

Composition in a piece of music is more important. The spectral description of such a signal is complex and varies with time in a peculiar manner depending on the type of music. Individual spectral components may vary in different manner. It is this aspect which makes modelling of each bin important. The order of the model for each bin has to be decided individually so that the variations are properly reflected in restored part.

2.MODEL AND MODEL ORDER ESTIMATION

2.1 Selection of model

Since many deterministic as well as stochastic discrete time processes in practice are well approximated by Rational Transfer Function Model, Transfer Function modelling is used as against extrapolation. System transfer function H(z) between input sequence (n_n) and output sequence (x_n) for Autoregressive Moving Average Model is given by

$$H(z) = B(z)/A(z)$$

where

A(z) = Z transform of autoregressive branch

$$= \sum_{m=0}^{p} a_m z^{-m}$$

B(z) = Z transform of moving average branch

$$= \sum_{m=0}^{q} b_m z^{-m}$$

Depending on whether none of the sets of parameters (a's & b's) vanish or either of them vanish, the process can be modelled as ARMA or only AR or only MA. However in general Autoregressive Model with proper order can describe the process properly although the process is close to ARMA or MA. Thus an autoregressive process of order p is expressed as

$$x_n = n_n - \sum_{k=1}^{p} a_k x_{n-k}$$

where the sequence x_n is a linear regression on itself and n_n representing the error. Present value of the process is expressed as weighted function of the past values and a noise term.

The Yule Walker equation

$$\begin{bmatrix} Rxx(0) & Rxx(-1) & \ldots\ldots\ldots & Rxx(-p) \\ Rxx(1) & Rxx(0) & \ldots\ldots\ldots & Rxx(-(p-1) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ Rxx(p) & Rxx(p-1) & \ldots\ldots\ldots & Rxx(0) \end{bmatrix} \begin{bmatrix} 1 \\ a1 \\ \cdot \\ \cdot \\ \cdot \\ ap \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

describes the linear relationship between the AR parameters and the autocorrelation function.

2.2 Selection of Model Order

In view of making model order selection adaptive in nature the following criteria were tried.

(1) FINAL PREDICTION ERROR (FPE)

This criterion is defined as

$$FPE = E_p \left(\frac{N+p+1}{N-p-1} \right)$$

The model order (p) selected is one for which FPE is minimum.

(2) AKAIKE INFORMATION CRITERION (AIC)

The AIC determines the model order by minimizing an information function

$$AIC = N \ln(E_p) + 2p$$

The model order (p) selected is one for which AIC is minimum.

(3) MINIMUM DESCRIPTION LENGTH (MDL)

According to MDL criterion, we compute

$$MDL = N \ln(E_p) + p \ln(N)$$

Again for the correct model order (p) MDL is minimum.

2.3 Modelling of Signal represented in frequency domain

In the present system, the signal being processed is recorded music. Composition of such a signal is quite complex since it would normally contain vocal music accompanied by musical instruments. In case of such a combination the spectrum representing signal not only has distinct frequencies/bands, the strengths of individual components vary in different manner. The signal is therefore represented in a better manner in frequency domain. Under these circumstances it would be proper to predict values of frequency components rather than signal values in time domain. This approach also simplifies the matrix operations involved in the modeling procedure since matrices to be dealt with for individual bins are smaller in size.

This processing involves transformation of signal in time domain to frequency domain through FFT. Proper Windowing and Overlapping of frames minimize distortion. Individual bin values over several frames are used for prediction. Predicted values of all the bins are transformed to time domain through IFFT.

3.IMPLEMENTATION

The system is implemented on a 386 based IBM compatible computing machine with fast ADC card. The signal is sampled at 44 Ksamples per second and through expanded memory is stored on the hard disc. Due to hardware limitation further processing is done in an offline manner. The basic technique, however does not need offline processing. The block diagram of the system implementation is shown in Fig.1. Stored samples are divided into overlapping frames and Fourier transform for these frames is evaluated with Kaiser window preceding. Individual bin values for N = 512 sample frame are evaluated for k = 15 consecutive overlapped frames and are used to predict bin values for (k+1) frame. IFFT is evaluated taking into consideration all such predicted bin values for (k+1) frame. Inverse windowing and selection of proper portion of computed IFFT for outputting is carried out in order to minimize distortion.

Important features of software are as follows. Expanded memory is used for larger storage. Kaiser window with β = 5 is selected as window function. Overlapping of frames is done through software while acquiring samples for FFT. The Cooley Tukey algorithm is used for FFT and Burg's algorithm is used for model order evaluation.

4.RESULTS

Fig. 2 & Fig. 3 shows results under different conditions.These were obtained by considering an artificial breakage introduced in the recording so that the comparison can be made between actual signal and the predicted signal during the breakage interval. Important features are tabulated in Fig. 4 and are as follows -

The prediction error decreases as the number of frames taken for prediction increases. The error is less when every bin is allowed to use an adaptive model order for itself. For most of the bins the model order selected was found to be between 4 and 9. Some other important observations are as follows -

For a rapidly changing piece of music (Fig. 4, Part A) the number of bins with model order between 7 - 9 is 242. The number of bins with model order between 10 - 12 and 13 - 15 is also large as compared to those for a slowly changing piece of music (Fig.4, Part B). As against this, for a slowly changing piece of music bins with model order between 4 - 6 is 256. Number of bins with model order between 1 - 3, which is 96, is also large as compared to those for a rapidly changing piece of music. Least error occurs when the model order selection is adaptive in nature. For a fixed order model (for each bin), the order if chosen according to the type of music, i.e. large (7) for faster and small (4) for slower, introduces lesser error. Selection of very high model order (14) introduces the largest error in both cases, more so with slowly changing piece of music. Amongst three optimization criteria used, it is observed that the Akaike Information Criterion (AIC) gives the least error in both the cases.

5.CONCLUSION

It may be concluded that making model of every bin adaptive gives least error. But this also requires largest time. A compromise may be made by having the model order fixed for all bins, with model order chosen intelligently as per the type music.

6.REFERENCES

(1) KAY AND MARPLE "Spectral Analysis",Proceedings of the IEEE, Vol. 69, No. 11, pp. 1381-1413, November 1981.

(2) STEPHEN KAY "Modern Spectral Estimation",Prentice Hall Edition.

(3) K.S.JOG AND AJAY INGLE "Sound Restoration in Damaged Acoustical Recording",Proceedings of the IEEE, Vol. 2, pp. 257-260, March 1992.

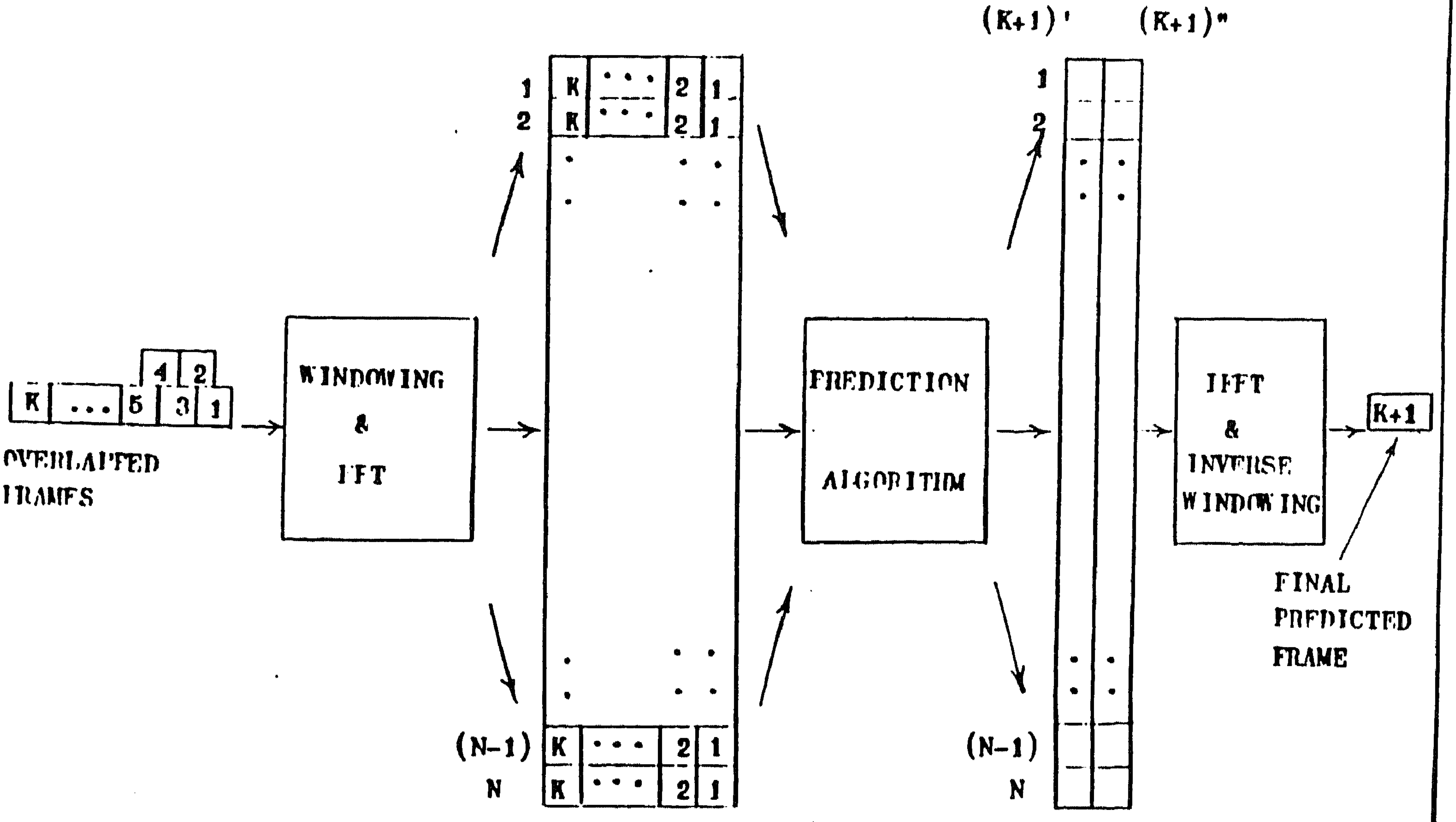


FIG.NO.:1 BLOCK DIAGRAM OF SYSTEM IMPLEMENTATION

RESULTS WITH A PIECE OF FAST MUSIC

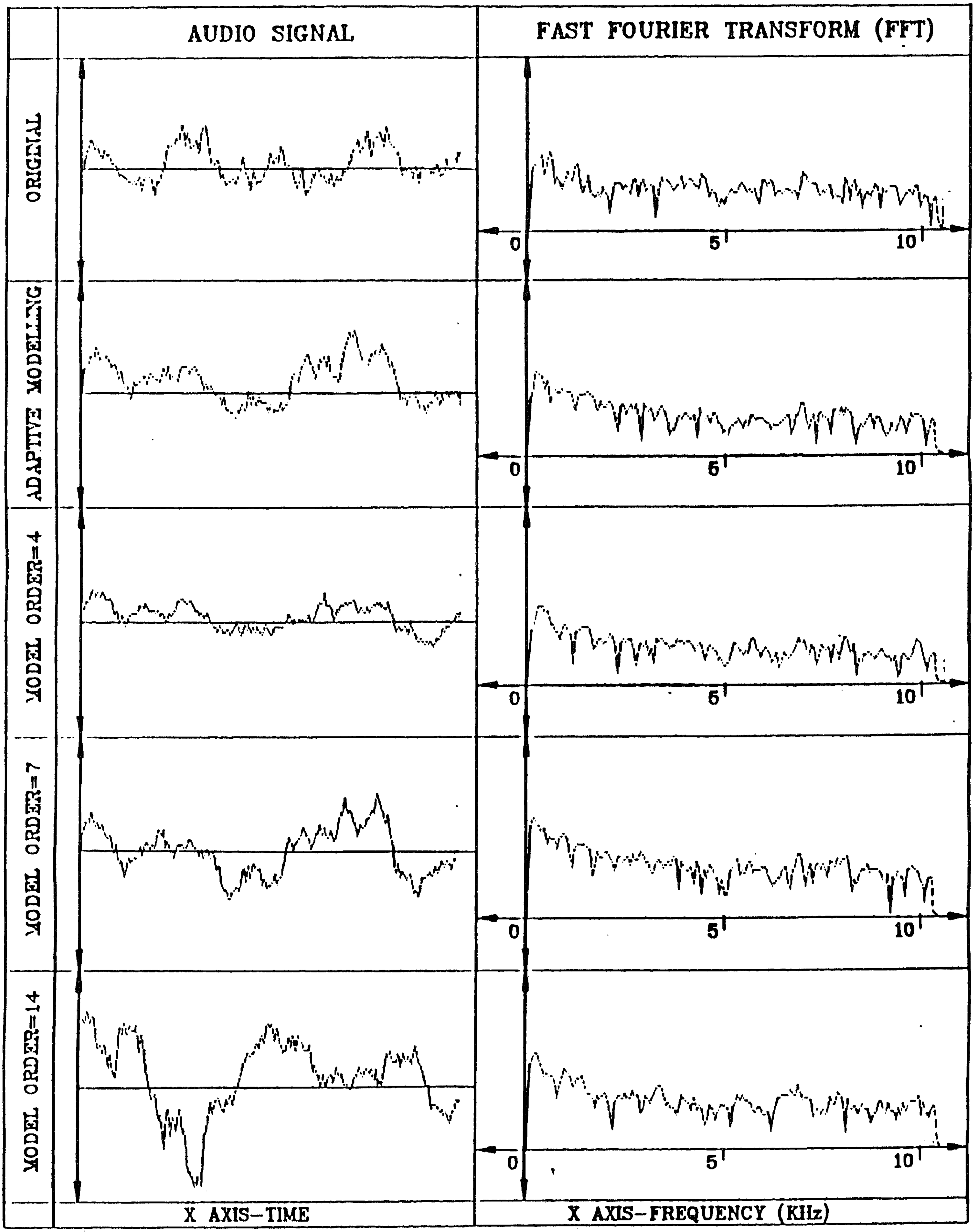


FIG.NO.: 2

RESULTS WITH A PIECE OF SLOW MUSIC

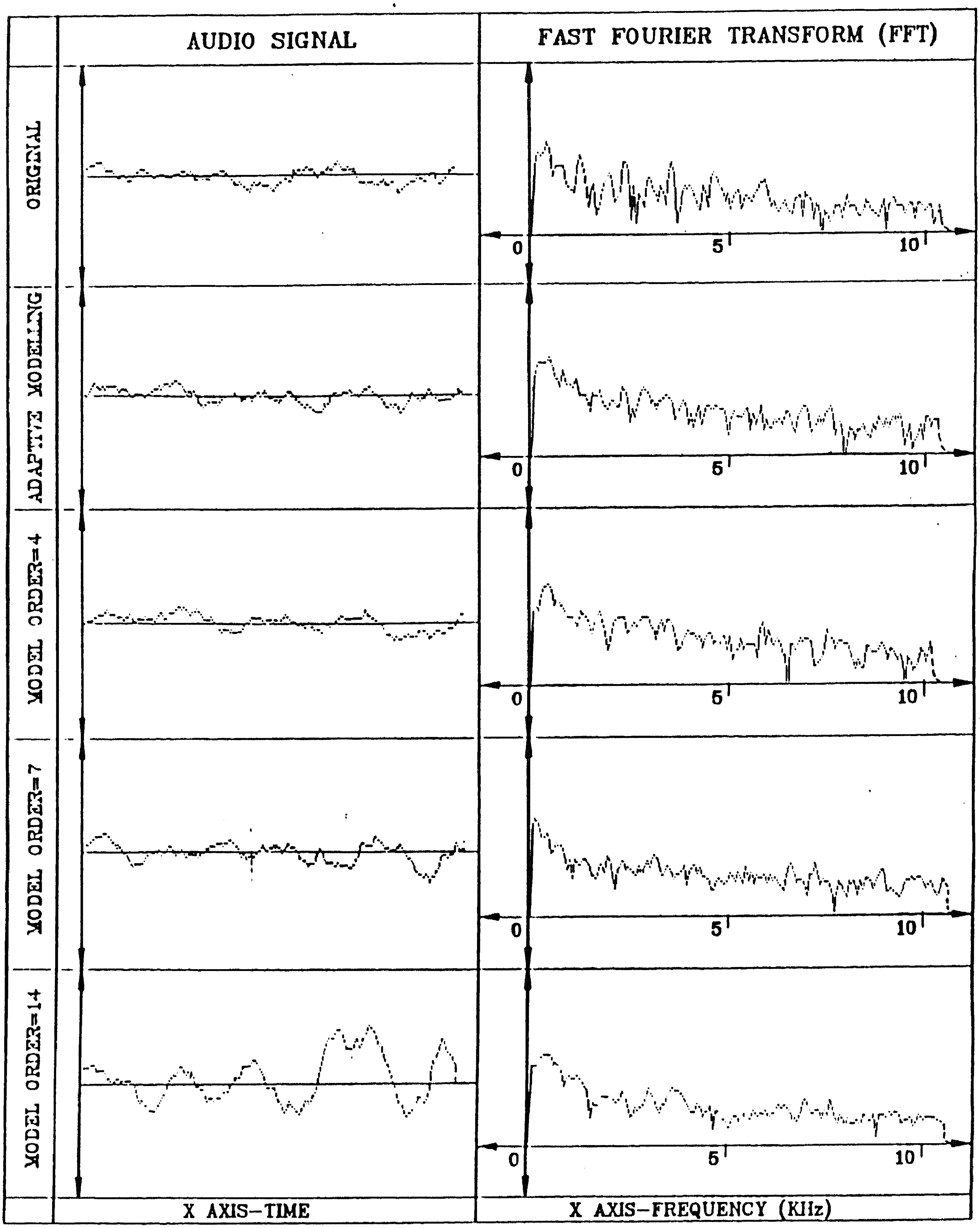


FIG.NO.: 3

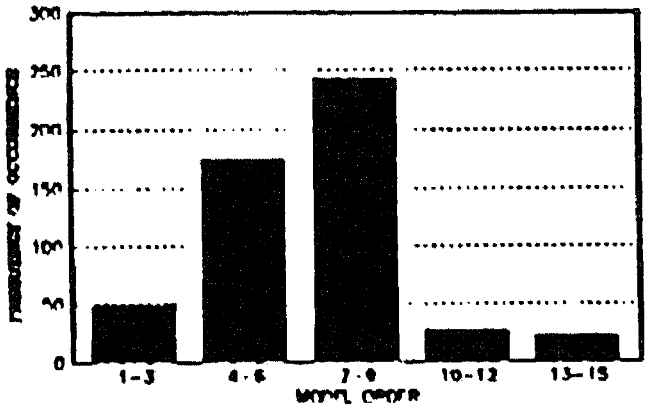
(A) RESULTS WITH A PIECE OF FAST MUSIC

	NATURE OF MODELLING USED	ERROR IN PREDICTION (%)	TIME TAKEN FOR PREDICTION (SEC)
1	Adaptive Modelling	4.1	75
2	Model Order = 4	7.4	12
3	Model Order = 7	5.3	15
4	Model Order = 14	10.9	40

ERROR INVOLVED IN PREDICTION WITH THE FOLLOWING CRITERIA :

1. Akaike Information Criteria (AIC) = 4.1 %
2. Final Prediction Error (FPE) = 5.07 %
3. Minimum Description Length (MDL) = 5.9 %

The Error involved when seven frames were taken for prediction with adaptive modelling was 6.2 %.
The Error involved when fifteen frames were taken for prediction with adaptive modelling was 4.1 %.
The bar chart gives the frequency of occurance of a particular model order when modelling is adaptive in nature.



(B) RESULTS WITH A PIECE OF SLOW MUSIC

	NATURE OF MODELLING USED	ERROR IN PREDICTION (%)	TIME TAKEN FOR PREDICTION (SEC)
1	Adaptive Modelling	3.6	75
2	Model Order = 4	4.8	12
3	Model Order = 7	7.3	15
4	Model Order = 14	11.2	40

ERROR INVOLVED IN PREDICTION WITH THE FOLLOWING CRITERIA :

1. Akaike Information Criteria (AIC) = 3.6 %
2. Final Prediction Error (FPE) = 4.4 %
3. Minimum Description Length (MDL) = 5.0 %

The Error involved when seven frames were taken for prediction with adaptive modelling was 5.3 %.
The Error involved when fifteen frames were taken for prediction with adaptive modelling was 3.6 %.
The bar chart gives the frequency of occurrence of a particular model order when modelling is adaptive in nature.

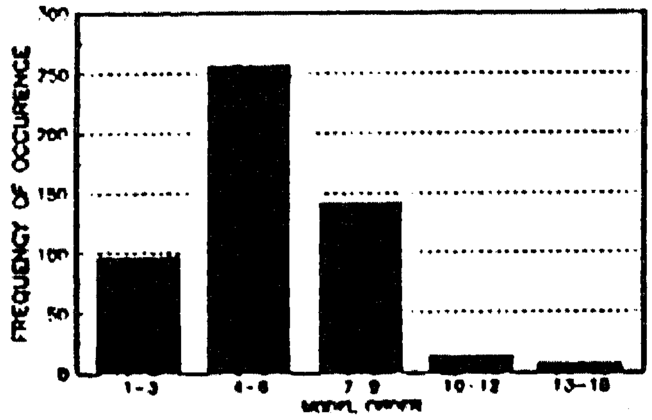


FIG. NO. : 4

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 111-134

ON THE GAUSSIAN APPROXIMATION ANALYSIS OF M-ARY SIGNALLING IN DIRECT DETECTION OPTICAL CDMA COMMUNICATION SYSTEMS

Ms.M. Meenakshi and Dr. V.C. Ravichandran

ABSTRACT

Performance of direct-detection Optical code division multiple-access communication systems with Avalanche photo diode (APD) photo detectors for binary and M-ary (M=4) signalling formats are analysed and presented in this paper. Analysis using Gaussian approximation for the symbol error probability (PSE) is carried out with APD's and arbitrary {0,1} - valued optical signature sequences. Multiple user interference, inter-symbol interference, shot noise and receiver thermal noise that cause variations in the symbol error probability are included in this analysis. Optical orthogonal codes with auto- and cross correlation coefficients of unity are considered for the numerical analysis. The results of the above analysis reveals that (i) the symbol error probability decreases for higher values of mean signal photon count per symbol, (ii) the symbol error probability offered by this system is minimum for an optimal range of the mean gain (G) of APD, (iii) as the number of users (K) increases the system performance is degraded, (iv) systems with Q-ary signalling scheme perform better than that of binary signalling scheme with respect to symbol error probability, nevertheless system complexity cannot be minimized.

Dr.V.C.Ravichandran, Asst. Professor, School of Electronics and Communication Engineering, College of Engineering, Anna University, Madras - 600 025.

Ms. M.Meenakshi, Junior Research Fellow, School of Electronics and Communication Engineering, College of Engineering, Anna University, Madras - 25.

I. INTRODUCTION

The CDMA techniques have been well applied to digital radio-frequency (RF) systems and good results obtained in the recent past. At present there is an increasing interest in applying this technique to optical communication systems to utilize the enormous bandwidth offered by modern fiber links.

Analysis is carried out in determining the symbol error probability with binary and M-ary signalling formats for multiple user systems. The random nature of the APD photo detector poses a major difficulty in the analysis. Due to the complicated nature of the statistical description of the response of the APD to incident photons, the APD output is modelled as a Gaussian process. The use of Gaussian approximation in the calculation of system error probabilities for binary signalling format has been shown [1] to be quite accurate in situations where the bulk leakage current of the APD itself is of the order of nA, the APD gain is a few hundred, and background light levels are not negligible.

The model for an optical CDMA communication system is shown in Fig.1. There are 'K' simultaneous users among 'U' multiple users transmitting data in binary or M-ary format, continuously and asynchronously. The optical characteristics of the 'U' lasers used are assumed to be identical. Also perfect synchronization between the transmitter and the intended receiver, is assumed. At the transmitter each user is assigned M mutually orthogonal {0,1}-valued periodic spreading waveforms, $a^{(k)}{}_j\ (t)$, $j = 1,2,\dots M$; $k = 1,2\dots K$ of period T_s. To transmit a data symbol, the kth user modulates its laser with the corresponding spreading waveform for a symbol duration of T_s seconds. The laser lights are assumed to be noncoherent so that the optical signals of the users are additive in intensity. At the receiving end, the received optical signal is a super-imposed sum of the 'K' delayed laser outputs. The photo detector converts the received optical signal into an electrical signal. The optimum receiver for this orthogonal signal set consists of a bank of M integrate and dump (or matched) filters as shown in Fig.2. This receiver correlates the received signal with the corresponding spreading waveform, $a^{(1)}_j$, $j = 1,2,3,\ldots M$, of the reference user (assumed to be the first user) and integrates for a data symbol duration of T_s seconds, to form the decision statistics Y_j, $j=1,2,\ldots M$ and decides that $a^{(1)}_j$ was present at its input during the corresponding signalling interval if

$$\max_i [Y_i] = Y_j; \; i = 1,2,\ldots M. \tag{1}$$

The front end pre-amplifier of the photo detector adds a thermal noise component to the decision statistics.

II. OPTICAL ORTHOGONAL CODE SEQUENCES

Direct detection optical CDMA systems require {0,1} valued code sequences with good correlation properties. Small cross correlation between the sequences are required for low multiple user interference (MUI), and small out-of-phase auto correlation are required for acquiring and maintaining synchronization between a transmitter and the desired receiver, in general.

An optical orthogonal code (OOC) is a family of {0,1} sequences with good auto- and cross-correlation properties. An $(N, \gamma, \lambda_a, \lambda_c)$ optical orthogonal code C is a family of {0,1} sequences of length N and weight γ which satisfy the following two properties.

(i) The auto- correlation property:

$$\sum_{t=0}^{N-1} a_t \cdot a_{t+\tau} \le \lambda_a \tag{2}$$

for any $a \in C$ and any interger τ, $0 < \tau < N$.

(ii) The cross-correlation property:

$$\sum_{t=0}^{N-1} a_t \cdot b_{t+\tau} \le \lambda_c \tag{3}$$

for any $a \neq b \in C$ and any integer τ.

A {0,1}- valued optical code sequence 'a' of length N and weight γ has an equivalent set representation given by $S_a = \{s_0, s_1, \ldots, s_{\gamma-1}\}$ [2] where s_i represents the position of the i th '1' in the sequence. The weight γ is the total member of 1s in 'a'. The difference rectangle, A_{ab} between sequence $a = (a_0, a_1, \ldots a_{N-1})$ and $b = (b_0, b_1, \ldots, b_{N-1})$, both with length N is defined as

$$A_{ab} = [s_{bj} - s_{ai}]; \quad \begin{array}{l} 0 \leq j \leq \gamma-1 \pmod N \\ 0 \leq i \leq \gamma-1 \pmod N \end{array} \qquad (4)$$

The integral entries in A_{ab} are the values of the index m such that the periodic correlation,

$$\theta_{ab} = \sum_{i=0}^{N-1} a_i \cdot b_{i+m}$$

are non zero. This difference rectangle method is very efficient in evaluating the auto- and cross-correlations [1].

In general, for a given code length N and weight γ, where $\gamma(\gamma-1) \leq N-1$ and $\lambda_a = \lambda_c = 1$, one can construct at most ϕ OOC's [2]. ϕ is called the size of the code. The value of ϕ is upper bounded by,

$$\phi \leq \left\lfloor \frac{N-1}{\gamma(\gamma-1)} \right\rfloor \qquad (5)$$

where the symbol $\lfloor x \rfloor$ denotes the integer portion of the real value x.

A total of $M \leq N$ mutually orthogonal sequences can be generated from a single sequence by using a shifted or a shifted reverse version of the sequence. An M-ary symbol can be transmitted using one of the M mutually orthogonal sequences. Using shifted reverse OOC sequences has an additional synchronization advantage since the periodic cross-corelation between an OOC and its reverse is bounded by 2 [1].

In the numerical analysis of this paper, one coincidence optical orthogonal codes are considered [2], that have periodic cross-correlations and out of phase periodic auto- correlations bounded by one. In this scheme, each user is assigned M mutually optical orthogonal code sequences of length N, with equal weights γ, for transmission. The spreading waveforms $a^{(k)}{}_j\ (t)$, $j = 1,2,3,\ldots M$, of the k th user can be expressed in terms of the code sequence $a^{(k)}{}_j = (a^{(k)}{}_{0j}, a^{(k)}{}_{1j}, a^{(k)}{}_{2j}, \ldots a^{(k)}{}_{(N-1)j},)$ as,

$$a^{(k)}{}_j\ (t) = \sum_{i=-\infty}^{\infty} a^{(k)}{}_{ij} \cdot P_{T_C}\ (t-iT_C) \qquad (6)$$

where $a^{(k)}{}_{(i+mN)j} = a^{(k)}{}_{ij}$ for $j = 1,2,\ldots M$ and for all 'i' and 'm'. Here $T_C = T_S/N$ is the chip time and $P_\tau(t)$ is a unit amplitude rectangular pulse of duration τ seconds defined by,

$$P_\tau(t) = \begin{cases} 1, & 0 < t < \tau \\ 0, & \text{otherwise} \end{cases} \qquad (7)$$

III. PHOTO DETECTORS

Avalanche photodiodes offer larger responsivity and are more sensitive to the incoming photons. Hence they are more widely used than PIN photodiodes. Optical intensity at the APD input results in the generation of primary photo electron hole pairs by the absorption of incident photons impinging onto the APD. The generation instants of these primary photo electron - hole pairs are modelled as a Poisson point process of rate $\lambda(t)$. Each primary pair then produces a random number of secondary electrons at the APD output by avalanche multiplication, which is assumed to be instantaneous. The number $\{g_1\}$ of secondary electrons produced by each primary, referred to as the APD gain, are assumed to be independant and identically distributed (i.i.d) random variables [3].

The mean value of the gain is $E\ [g] \equiv G$ and the mean square value of g is

$$v_g \equiv E\ [g^2] = [k_{eff}\ \ G + (2 - (1/G))\ (1 - k_{eff}]\ G^2 = FG^2 \qquad (8)$$

where $F = [k_{eff} G + (2 - (1/G)) (1 - k_{eff})]$ is the APD excess noise factor. In the above expression, k_{eff} (the ratio between the ionization probabilities per unit length for holes and electrons in the depletion region) is a fabrication - dependant parameter; the lower it is, the better.

IV. GAUSSIAN APPROXIMATION

The following analysis is done for the Q-ary (M=4) signalling scheme. In this section, the probability of symbol error PSE is evaluated by assuming that the decision statistics Y_j, $j = 1,2,3,4$, is approximately Gaussian. Under this assumption, only the mean and the various of Y_j need to be evaluated and PSE is expressed in terms of simple signal to noise ratios.

Referring to Fig.1., the modulating waveform for the kth laser can be written as,

$$\begin{aligned} d^{(k)}(t) = \ & [1-b^{(k)}(t)] \cdot [1-b^{(k)}(t+ T_s/2)] \cdot a^{(k)}{}_1(t) \\ & + [1-b^{(k)}(t)] \cdot [b^{(k)}(t+ T_s/2)] \cdot a^{(k)}{}_2(t) \\ & + [b^{(k)}(t)] \cdot [1-b^{(k)}(t+ T_s/2)] \cdot a^{(k)}{}_3(t) \\ & + [b^{(k)}(t)] \cdot [b^{(k)}(t+ T_s/2)] \cdot a^{(k)}{}_4(t) \end{aligned} \qquad (9)$$

where,

$$\begin{aligned} b^{(k)}(t) = \ & \sum_j \{b^{(k)}{}_{j0} \cdot P_T(t - (2j + 1) T_s/2) \\ & + b^{(k)}{}_{j1} \cdot P_T(t - (2j + 2) T_s/2)\} \end{aligned} \qquad (10)$$

Here $b^{(k)}(t)$ is the data waveform of the kth user and $b^{(k)}{}_{j0}$, $b^{(k)}{}_{j1} \epsilon (0,1)$ are the alphabets. $P_T(t)$ is defined by equation (7) and $a^{(k)}{}_j(t)$, $j = 1,2,3,4$ are the spreading waveforms.

The following discussion is phrased in terms of making a decision about the value of the data symbol $a^{(1)}{}_j$, $j = 1,2,3,4$, transmitted in the time interval $(0,T_s)$ by the first user. The generation instants of the primary photo electron hole pairs, $\{t_j\}$, conditioned on the relative time delays τ_k and the data symbols of all 'K' simultaneous users is modelled as a Poisson point process with inhomogenous rate,

$$\lambda(t) = \sum_{k=1}^{K} \lambda^{(k)} (t-\tau_k) + \lambda_{bl}; \tau_1 = 0 \tag{11}$$

Here λ_{bl}, is the APD bulk leakage current photon rate and $\lambda^{(k)}$ (t) is the photon rate from the kth user given by,

$$\lambda^{(k)} (t) = \lambda_s (1 - 1/\alpha) \; d^{(k)} (t) + \lambda_s/\alpha \tag{12}$$

In the above expression, $\lambda_s = \eta \; P_s/h\varphi$ is the signal photon rate with 'η' the APD quantum efficiency, 'P_s' the laser power in the ON state, '$h\varphi$' the energy of a photon and 'α' the laser modulation - extinction power ratio, which accounts for the laser leakage effect when a laser is in the OFF state. By combining the laser leakage photon rate and APD bulk leakage photon rate together to form a total background photon rate λ_o , equation (11) can be written as,

$$\lambda (t) = \sum_{k=1}^{K} \lambda'_s \; d^{(k)} (t - \tau_k) + \lambda_o \tag{13}$$

where $\lambda_o = \lambda_b + \lambda_{bl}$ with $\lambda_b = K \lambda_s/\alpha$, $\lambda'_s = \lambda_s(1-1/\alpha)$.

The APD surface leakage current is not multiplied by the APD gain [3]. We model the generation times $\{t_i\}$ of the APD surface leakage photon as an independent Poisson point process of constant rate λ_{sl}.

In Fig.2., the linear time invariant filter output, in units of number of electrons, is,

$$y_j (t) = 1/e \; [X (t) * h_s (t) + X_{sl} (t) * h_d (t) + n (t)] * h_j (t) \qquad j = 1,2,3,4$$

$$= \frac{1}{e} \left(\sum_{q=1}^{N(t)} e \; g_q \; \delta (t - t_q) \right.$$

$$+ \sum_{p=1}^{N_{s1}(t)} e\,\delta\,(t - t_p) \;+ n\,(t) \Bigg) * h_j\,(t)$$

$$j = 1,2,3,4$$

$$= \sum_{q=1}^{N(t)} g_q\, h_j\,(t - t_q) + \sum_{p=1}^{N_{s1}(t)} h_j\,(t-t_p)$$

$$+ \frac{1}{e} \int_{-\infty}^{t} n\,(\tau)\, h_j\,(t - \tau)\, d\tau; \; j = 1,2,3,4. \qquad (14)$$

where '*' denotes convolution, and $N(t)$ and $N_{s1}(t)$ are independent Poisson counting processes [4] with intensity functions $\lambda\,(t)$ and λ_{s1} respectively. For simplicity the APD filter responses have been chosen to be $h_s(t) = eg\delta\,(t)$ and $h_d\,(t) = e\delta\,(t)$. The linear filter impulse response is,

$$h_j\,(t) = a^{(1)}{}_j\,(T_s - t)\,\cdot P_{T_s}\,(t),\; j = 1,2,3,4. \qquad (15)$$

At time $t = T_s$, the decision statistics can be written as,

$$Y_j = y_j\,(T_s) = Z_j + Z_{s1} + W; \qquad j = 1,2,3,4. \qquad (16)$$

The noise $n(t)$ at the pre-amplifier input is modelled as additive white Gaussian noise with two sided power spectral density $2K_B\,T_e/R$ where K_B is the Boltzmann's constant, T_e is the equivalent noise temperature, and R is the APD load resistor. The random variable W in equation (16) is then zero mean Gaussian with variance equal to,

$$\sigma^2{}_W = \frac{2K_B T_e \gamma T_c}{e^2\,R} \qquad (17)$$

The APD response time must be fast enough to follow the spreading waveforms. One common design criterian is to set $RC_{APD} = T_c$ where C_{APD} is the APD junction capacitance.

In the receiver of Fig.2, the APD output is separately multiplied by $a^{(1)}{}_1\,(t)$, $a^{(1)}{}_2\,(t)$, $a^{(1)}{}_3\,(t)$, and $a^{(1)}{}_4\,(t)$ and integrated for the interval $(0, T_s)$ to form the sum Z_1, Z_2, Z_3 and Z_4 respectively, which represent the corresponding number of secondary electrons collected in the $a^{(1)}{}_1$, $a^{(1)}{}_2$, $a^{(1)}{}_3$, $a^{(1)}{}_4$, windows. Therefore Z_j, $j = 1,2,3,4$ can be written as follows.

$$Z_j = \begin{cases} S_j + I_{kj} + I_{1j} & ;\ \text{if } a^{(1)}{}_j \text{ was received} \\ I_{kj} + I_{1j} & ;\ \text{if } a^{(1)}{}_i,\ i \neq j \text{ was received} \end{cases} \qquad i,\ j = 1,2,3,4. \tag{18}$$

In the above expression S_j is the number of secondary electrons collected in the $a^{(1)}{}_j$ window due to the transmission of the sequence $a^{(1)}{}_j$, through laser 1.

The multiple user interference term I_{kj} is given by

$$I_{kj} = \sum_{k=0,k\neq 1}^{K} A_{kj} \tag{19}$$

where A_{kj} is the number of secondary electrons collected in the $a^{(1)}{}_j$ window due to radiation from laser 'k'.

The inter symbol interference term I_{1j} is given by,

$$I_{1j} = \sum_{i=1,i\neq j}^{4} A_{1i} \tag{20}$$

where A_{1i} is the number of secondary electrons collected in the $a^{(1)}{}_j$ window due to the transmission of the sequence $a^{(1)}{}_i$, $i = 1,2,3,4;\ i \neq j$; through laser 1.

A_{0j} corresponds to the secondary electrons collected in $a^{(1)}{}_j$ window due to the background radiation term. In more detail,

$$A_{kj} = \sum_{l=1}^{M_{kj}} g_l \tag{21}$$

and

$$A_{1i} = \sum_{l=1}^{M_{1i}} g_l \qquad ;i,j = 1,2,3,4;\ i \neq j \tag{22}$$

where M_{kj} is the number of primary electron-hole pairs that are generated during the $a^{(1)}{}_j$ window due to optical radiation from the kth laser, M_{1j} is the number of primary electron-hole pairs that are generated due to the signal $a^{(1)}{}_i$, $i \neq j$, $i=1,2,3,4$ being transmitted through laser 1, and g_l is the APD gain for the lth primary electron-hole pair.

For $K \geq 2$, the conditional means of the conditionally Poisson random variables M_{kj} are,

$$E[M_{kj}/\tau_k, d^{(k)}{}_{-1}, d^{(k)}{}_0] = \int_0^{T_s} \lambda'_s \, d^{(k)} \, (t - \tau_k) \, a^{(1)}{}_j \, (t) \, dt = \Lambda_{kj} \qquad (23)$$

Similarly,

$$E[M_{1i}/\tau_1, d^{(1)}{}_{-1}, d^{(1)}{}_0] = \int_0^{T_s} \lambda'_s \, d^{(1)} \, (t - \tau_1) \, a^{(1)}{}_j \, (t) \, dt = \Lambda_{1i}$$

$$i,j = 1,2,3,4; \; i \neq j \qquad (24)$$

Further more, $S_j = A_{1j}$ with $\Lambda_{1j} = E[M_{1j}] = \lambda'_s \, \gamma \, T_c$ and $\Lambda_{0j} = E[M_{0j}] = \lambda_o \, \gamma \, T_c$.

The discrete aperiodic cross-correlation between the code sequences $a^{(x)}$ and $a^{(y)}$ is defined as [5] [6],

$$c^{x,y} \, (l) = \begin{cases} \sum_{t=0}^{N-1-l} a^{(x)} \, (t) \cdot a^{(y)} \, (t + l) & , \; 0 \leq l \leq N-1 \\ \sum_{t=0}^{N-1+l} a^{(x)} \, (t-l) \cdot a^{(y)} \, (t) & , \; N-1 \leq l \leq 0 \\ 0 & , \; \text{Otherwise} \end{cases} \qquad (25)$$

A two user correlation parameter P for the sequences $a^{(x)}$ and $a^{(y)}$ (with the first user as the reference user), is defined as

$$p^{x,y} \, (l) = c^{x,1} \, (l-N) + c^{y,1} \, (l); \; l = 0, \ldots N \qquad (26)$$

The conditional means Λ_{kj} can be expressed in terms of the P parameters as [1],

$$\Lambda_{kj}\,(\tau_k,\ d^{(k)}_{-1},\ d^{(k)}_{0}) = \lambda'_s\ [P^{x,y}\ (l)\ .\ T_c$$

$$+\ \{\ P^{x,y}\ (l+1) - P^{x,y}\ (l)\}\ (\tau_k - lT_c)]$$

$$\text{for } 0 \le l \le \tau_k \le (l+1)\ T_c \le NT_c \tag{27}$$

The superscripts x and y, depend on two consecutive data symbols arriving at the receiver. They are given by,

$$x = \begin{array}{l} k_1,\ d^{(k)}_{-1} = 00 \\ k_2,\ d^{(k)}_{-1} = 01 \\ k_3,\ d^{(k)}_{-1} = 10 \\ k_4,\ d^{(k)}_{-1} = 11 \end{array} \qquad y = \begin{array}{l} k_1,\ d^{(k)}_{0} = 00 \\ k_2,\ d^{(k)}_{0} = 01 \\ k_3,\ d^{(k)}_{0} = 10 \\ k_4,\ d^{(k)}_{0} = 11 \end{array} \tag{28}$$

where k_j in general denotes the jth symbol of the kth user.

A. Mean value of the output

It follows from equations (16) and (18) that,

$$E[Y_j/a^{(1)}_j \text{ transmitted}] = E[S_j] + E[I_{kj}] + E[I_{1j}] + E[Z_{s1}] \tag{29}$$

Mean value of the signal term is given by,

$$E[S_j] = G.E[M_{1j}] = G\lambda'_s\ \gamma\ T_c \tag{30}$$

Mean value of the surface leakage component is given by,

$$E[Z_{s1}] = \lambda_{s1}\ \gamma\ T_c \tag{31}$$

Mean value of the multiple user interference together with that of the back ground radiation is given by,

$$E[I_{kj}] = E\ [\sum_{k=0,k\neq 1}^{K} A_{kj}] = G\ \lambda_o\ \gamma\ T_c + \sum_{k=2}^{K} G.E[\Lambda_{kj}] \tag{32}$$

Delays τ_k are uniform in $(0,T_s)$ and data symbols are assumed to be equally likely. Therefore for $K \ge 2$,

$$E[\Lambda_{kj}] = \lambda'_s \; T_c \; \gamma^2/N \qquad (33)$$

Thus from equations (19), (32) and (33), it follows that,

$$E[I_{kj}] = G\lambda_o \; \gamma \; T_c + (K-1) \; G \; \lambda'_s \; T_c \; \gamma^2/N \qquad (34)$$

Similarly from equation (20), the mean of the inter symbol interference parameter can be obtained as

$$E[I_{1j}] = G \; E[\Lambda_{1i}] = G \; \lambda'_s \; T_c \; \gamma^2/N \qquad (35)$$

Therefore,

$$\begin{aligned} E[Y_j/a^{(1)}{}_j \text{ transmitted}] &= m_{Yj/a(1)j} \\ &= G \; \lambda'_s \; \gamma \; T_c + \lambda_{s1} \; \gamma \; T_c + G \; \lambda_o \; \gamma \; T_c + K \; G\lambda'_s \; T_c \; 2\gamma/N \\ &= G \; \mu_s + (\lambda_{s1}/\lambda'_s) \; \mu_s + G \; (\lambda_o/\lambda'_s) \; \mu_s + G \; \mu_s \; q \; K \qquad (36) \end{aligned}$$

where $\mu_s = \lambda'_s \; T_c \; \gamma$ is the mean signal photo-electron count per symbol, and $q = \gamma/N$ is the average duty cycle of the code sequence. It follows from equations (18) and (36) that,

$$\begin{aligned} E[Y_j/a^{(1)}{}_i, \; i \neq j, \text{transmitted}] &= m_{Yj/a(1)i} \\ &= (\lambda_{s1}/\lambda'_s) \; \mu_s + G \; (\lambda_o/\lambda'_s) \; \mu_s + G \; \mu_s \; q \; K \qquad (37) \end{aligned}$$

B. Variance of the Output

The variance of the signal term is given by,

$$\sigma^2{}_{sj} = v_g \; \lambda'_s \; \gamma \; T_c \qquad (38)$$

where $v_g = FG^2$, is the second moment of the APD gain.

The variance of the surface leakage current component is given by,

$$\sigma^2{}_{Zs1} = \lambda_{s1} \; \gamma \; T_c \qquad (39)$$

Therefore,

$$\sigma^2_{Y_j} \,/\, a^{(1)}_j = v_g \, \lambda'_s \, \gamma \, T_c + \lambda_{s1} \, \gamma \, T_c + \frac{2 K_B T_e}{e^2 R} \, \gamma \, T_c + \sigma^2_{I_{kj}} + \sigma^2_{I_{1j}} \qquad (40)$$

The variance of the multiple user interference term can be evaluated as follows.

$$\sigma^2_{I_{kj}} = \sum_{k=0,\ k\neq 1}^{K} \left[E\,[A^2_{kj}] - E[A_{kj}]^2 \right] \qquad (41)$$

$$E[A^2_{kj}] = \frac{v_g \, \lambda'_s \, \gamma^2 \, T_c}{N} + G^2 \, E[\Lambda^2_{kj}] \qquad (42)$$

Using equation (27), it can be shown that,

$$E[\Lambda^2_{kj}] = \frac{1}{16 \, T_s} \sum_{x,y \, \epsilon \{k_1,k_2,k_3,k_4\}} r^{x,y} \qquad (43)$$

where,

$$r^{x,y} = \frac{\lambda'^2_s \, T_c^3}{3} \sum_{l=0}^{N-1} \left[p^{x,y}(l+1)^2 + p^{x,y}(l)^2 + p^{x,y}(l+1) \, p^{x,y}(l) \right] \qquad (44)$$

Considering the variance of the background radiation term we have,

$$E[A^2_{0j}] - E[A_{0j}]^2 = v_g \, \lambda_o \, \gamma \, T_c \qquad (45)$$

Therefore we have,

$$\sigma^2_{I_{kj}} = v_g \, \lambda_o \, \gamma \, T_c + v_g \, \frac{\lambda'_s \, \gamma^2 \, T_c}{N} (K-1)$$

$$\frac{G^2 \, \lambda'^2_s \, T^2_c}{48 \, N} \sum_{k=2}^{K} \sum_{x,y \, \epsilon \, \{k_1,k_2,k_3,k_4\}} r^{x,y}$$

$$\frac{G^2 \lambda'^2_s T^2_c \gamma^4}{N^2} (K-1) \tag{46}$$

Similarly the variance of the inter symbol interference term is evaluated,

$$\sigma^2_{I1j} = \frac{v_g \lambda'_s \gamma^2 T_c}{N} + \frac{G^2 \lambda'^2_s T^2_c}{27 N} \sum_{p,q \in \{a^{(1)}_2, a^{(1)}_3, a^{(1)}_4\}} i^{p,q}$$

$$- \frac{G^2 \lambda'^2_s T^2_c \gamma^4}{N^2} \tag{47}$$

where

$$i^{p,q} = \sum_{l=0}^{N-1} [p^{p,q} (l+1)^2 + p^{p,q}(l)^2 + p^{p,q}(l+1) \; p^{p,q}(l)] \tag{48}$$

Therefore the conditional variance of Y_j becomes,

$$\sigma^2_{Y_j/a^{(1)}_j} = v_g \lambda'_s \gamma T_c + \lambda_{s1} \gamma T_c + \frac{2 K_B T_e \gamma T_c}{e^2 R}$$

$$+ v_g \lambda_o \gamma T_c$$

$$+ \frac{K v_g \lambda'_s \gamma^2 T_c}{N} - \frac{G^2 \lambda'^2_s T_c^2 \gamma^4 K}{N^2}$$

$$+ \frac{G^2 \lambda'^2_s T_c^2}{3N} \left[\frac{1}{16} \sum_{k=2}^{K} \sum_{x,y \in \{k_1, k_2, k_3, k_4\}} r^{x,y} \right.$$

$$\left. + (1/9) \sum_{p,q \in \{a^{(1)}_2, a^{(1)}_3, a^{(1)}_4\}} i^{p,q} \right]$$

$$= v_g \mu_s + \frac{\lambda_{s1} \mu_s}{\lambda'_s} + \frac{2K_B T_e \mu_s}{e^2 R \lambda'_s} + \frac{v_g \lambda_o \mu_s}{\lambda'_s}$$

$$+ \quad K\, v_g\, \mu_s q - G^2 \mu_s^2\, K\, q^2$$

$$+ \quad \frac{G^2 \mu_s^2}{3N\, \gamma^2} \left[\frac{1}{16} \sum_{k=2}^{K} \sum_{x,y\, \epsilon \{k_1,\, k_2,\, k_3,\, k_4\}} r^{x,y} \right.$$

$$\left. + \quad (1/9) \sum_{p,q\, \epsilon \{a^{(1)}_2,\, a^{(1)}_3,\, a^{(1)}_4\}} i^{p,q} \right] \tag{49}$$

Similarly, the conditional variance of y_j, when $a(1)_i$, $j \neq i$ is transmitted is given by,

$$\begin{aligned} \sigma^2_{Yj/a^{(1)}i,\ i \neq j} &= \sigma^2_{Y_j/a^{(1)}j} - \sigma^2_{s_j} \\ &= \sigma^2_{Y_j/a^{(1)}j} - v_g\, \lambda_s'\, \gamma T_c \end{aligned} \tag{50}$$

C. Symbol Error Probablity

Finally the probability of symbol error assuming $a^{(1)}_1$ as the transmitted symbol, is approximated as,

$$\text{PSE} = 1/4 \left\{ Q \left[\frac{m_{Y\ /a(1)_1} - \theta}{\sigma_{Y_1/a(1)1}} \right] + Q \left[\frac{\theta - m_{Y_1/a(1)2}}{\sigma_{Y_1/a(1)2}} \right] \right.$$

$$\left. + \quad Q \left[\frac{\theta - m_{Y_1/a(1)3}}{\sigma_{Y_1/a(1)3}} \right] + Q \left[\frac{\theta - m_{Y1/a(1)4}}{\sigma_{Y_1/a(1)4}} \right] \right\} \tag{51}$$

In general for any $M = 2^n$ we can write,

$$\text{PSE} = 1/M \left\{ Q \left[\frac{m_{Y_1/a(1)_1} - \theta}{\sigma_{Y_1/a(1)1}} \right] + (M-1)\, Q \left[\frac{\theta - m_{Y_1/a(1)i}}{\sigma_{Y_1/a(1)i}} \right] \right\} \tag{52}$$

where $i \neq 1$,

since, the values of $m_{Y_1/a(1)_i}$ and $\sigma^2_{Y_1/a(1)_i}$ are approximately the same for all $i \leq M$ except $i = 1$ in this case.

The optimum threshold 'θ' is obtained here, by equating

$$\left[\frac{m_{Y_1/a(1)_1} - \theta}{\sigma_{Y_1/a(1)_1}} \right] \quad \text{and} \quad \left[\frac{\theta - m_{Y_1/a(1)_i}}{\sigma_{Y_1/a(1)_i}} \right]_{i \neq 1}$$

Thus,

$$\theta = \frac{\sigma_{Y_1/a(1)_1} \cdot m_{Y_1/a(1)_i} + \sigma_{Y_1/a(1)_i} \cdot m_{Y_1/a(1)_1}}{\sigma_{Y_1/a(1)_1} + \sigma_{Y_1/a(1)_i}} \tag{53}$$

By assuming that the threshold 'θ' lies in the range $\theta_1 \leq \theta \leq \theta_2$, the symbol error probability can be approximated by,

$$PSE = 1/M \left\{ Q \left[\frac{m_{Y_1/a(1)_1} - \theta_1}{\sigma_{Y_1/a(1)_1}} \right] + (M-1)\, Q \left[\frac{\theta_2 - m_{Y_1/a(1)_i}}{\sigma_{Y_1/a(1)_i}} \right] \right\}. \tag{54}$$

where $i \neq 1$, $i \leq M$

and

$$\theta_1 = \frac{\theta \cdot \sigma_{Y_1/a(1)_1}}{m_{Y_1/a(1)_1}} \tag{55}$$

and

$$\theta_2 = \frac{\theta \cdot \sigma_{Y_1/a(1)_i}}{m_{Y_1/a(1)_i}} \qquad i \leq M,\ i \neq M \tag{56}$$

V. NUMERICAL RESULTS

The following parameter values have been assumed for the numerical analysis.

k_{eff} = effective ionization ratio = 0.01

R = APD load resistance = 1030 Ω

T_e = Preamplifier equivalent noise temperature = 1100°K

η = APD quantum efficiency = 75%

$$\frac{No}{2} = \frac{2\ K_B T_e}{e^2\ R} = \text{thermal noise two - sided PSD}$$

$$= 1.15 \times 10^{15}\ J/\Omega/C^2$$

λ_{bl} = APD bulk leakage current photon rate = 6.12×10^5 photons/sec

λ_{sl} = APD surface leakage current photon rate = 7.46×10^{10} photons/sec.

G = APD mean gain = 100

α = Laser - modulation extincton power ratio = ∞.

The received optical power is assumed to be 50 nw at a wavelength of 840 nm. Optical orthogonal codes with auto and cross correlation coefficient values λ_a and λ_c equal to unity are used. Sequences with length N = 341 and weight γ = 5 is used in this analysis. The effect of the mean gain of the APD on the probability of symbol error is observed, in both binary and Q-ary signalling schemes and are shown in Fig. 3. It is seen that for just one user (K = 1), the value of PSE in both cases is of the order of 10^{-30} for the values of 'G' close to 100. As the number of users 'K' in the system increases, the mean gain G is required to be a few hundreds to achieve minimum symbol error probability.

To study the effect of the mean signal photon count per symbol 'μ_s' on the system performance is a mean gain of 100 is assumed. The comparison of the performance of binary and Q-ary signalling schemes with respect to the mean signal photon count per symbol 'μ_s' is shown n Fig. 4. By considering the symbol error probability as the criterion for system performance, Q-ary systems perform better compared to binary systems. However as the number of users increases the symbol error probability of the Q-ary system approaches that of the binary system. The system performance improves as the mean signal photon count per symbol increases.

It is observed from Fig. 5 that for a code length of 341 and a PSE of 10^{-10}, the maximum number of simultaneous users 'K' is 5 for all values of 'μ_s'. Assuming that only 20% of the users in the CDMA network are transmitting at any instant of time a maximum number of 25 users can be accommodated in this network. By increasing the length of the code sequence used $N > 341$ the maximum number of users can be increased.

VI. CONCLUSION

Performance of an optical CDMA system using Avalanche photodiode photodetector for binary and M-ary signalling scheme are analyzed with Gaussian approximation for the symbol error probability.

Based on the above analysis carried out, the following inferences are made.

i. The probability of symbol error (PSE) decreases for higher values of the mean signal photon count per symbol.

ii. The probability of symbol error (PSE) offered by this system is minimum for an optimal range of the mean gain of APD
   a) 100 to 120 for values of $K = 1$ and 2
   b) 160 to 200 for values of $K = 5$

iii. As the number of users increases, $K > 5$ the system performance is degraded.

iv. Systems with Q-ary ($M=4$) signalling scheme perform better than binary signalling scheme with respect to the symbol error probability, nevertheless the system complexity cannot be minimized.

The above analysis is extended for higher values of M i.e. M = 8, 16, 32, 64, for all values of K = 1,2,5,10,16, and will be presented later.

REFERENCES

[1] A.W.Lam and A.M. Hussain,"Performance Analysis of Direct Detection Optical CDMA communication systems with Avalanche Photodiodes" IEEE. Trans. Commun., Vol. 40, No.4, pp. 810-820, April 1992.

[2] F.R.K.Chung, J.A. Salehi and V.K. Wei, "Optical Orthogonal codes : Design, analysis and applications", IEEE Trans. Inform. Theory, Vol. IT-35, pp. 595-604, May 1989.

[3] F.M. Davidson and X. Sun, "Gaussian approximation versus nearly exact performance analysis of optical communication systems with PPM signalling and APD receivers", IEEE Trans. Commun., Vol. 36, pp. 1185-1192, Nov. 1988.

[4] D.L Synder, Random Point Processes, New York : Wiley, 1975.

[5] M.B. Pursley, "Performance evaluation for phase coded spread-spectrum multiple user communication - Part I : System Analysis", IEEE Trans. Commun., Vol. COM-25, pp. 795-803, Aug. 1977.

[6] D.V. Sarwate, and M.B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," Proc. IEEE, Vol. 68, pp. 593-619, May 1980.

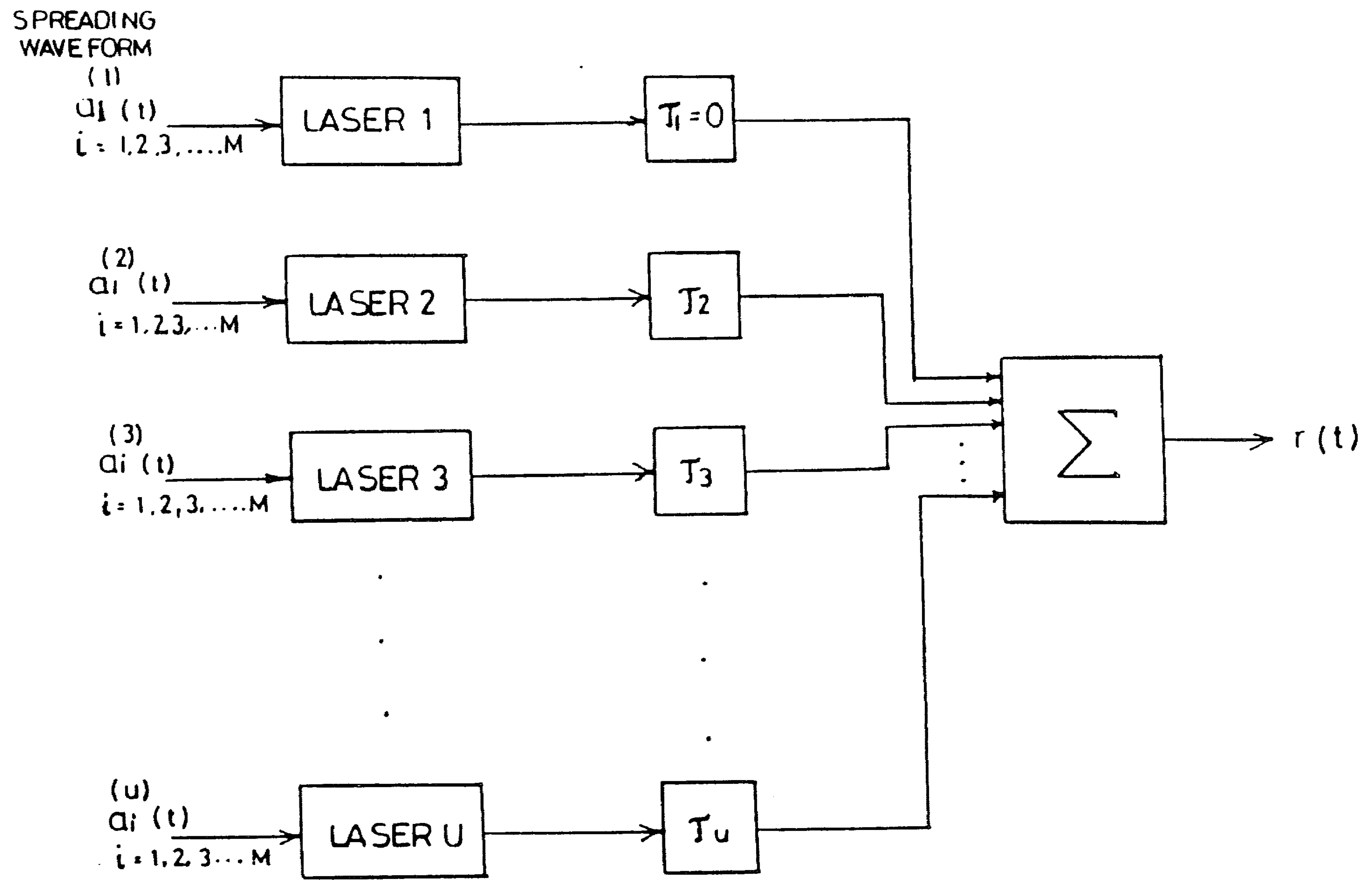


FIG 1 OPTICAL CDMA SYSTEM - TRANSMITTER MODEL

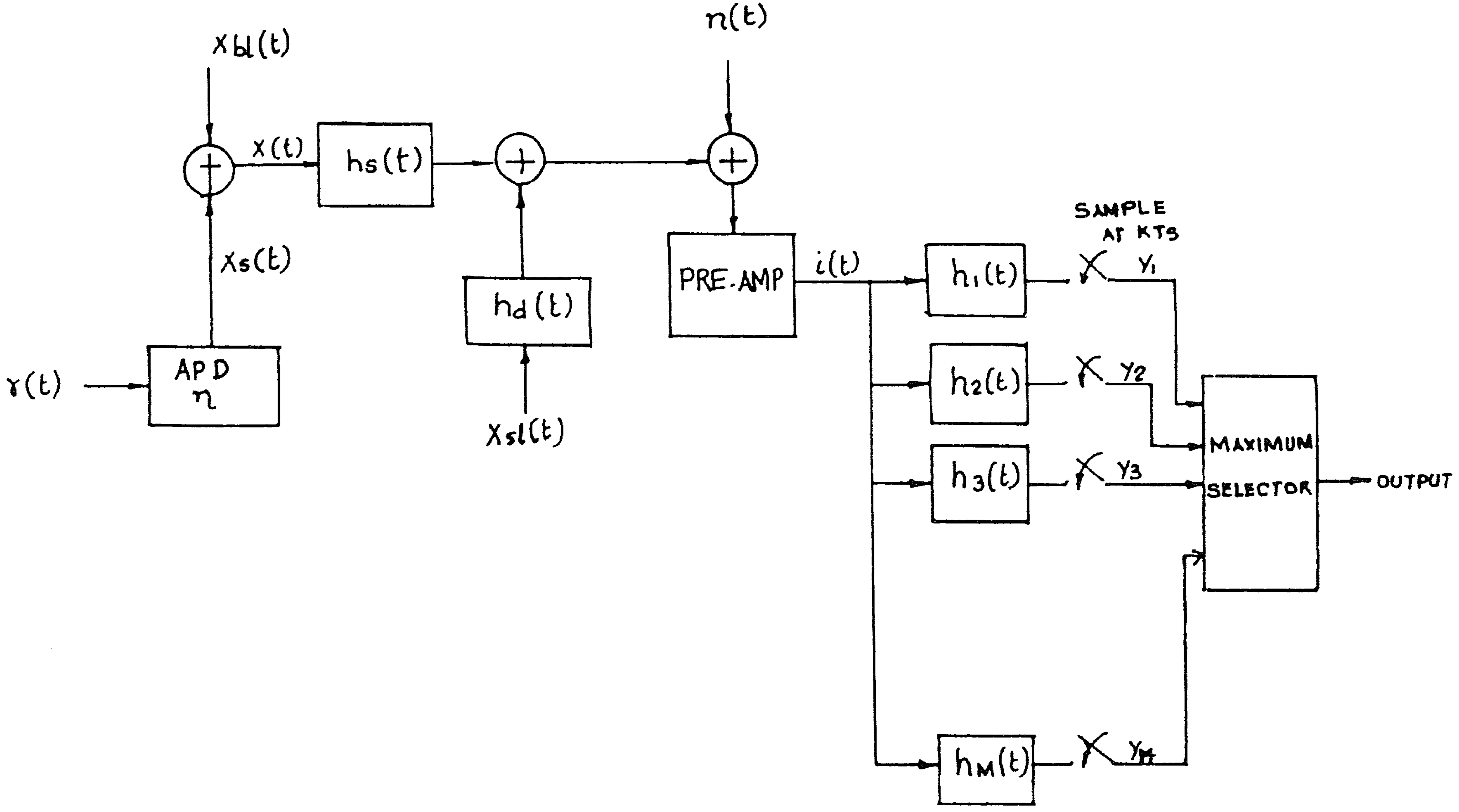


FIG2 OPTICAL CDMA SYSTEM RECEIVER MODEL

GATEWAY

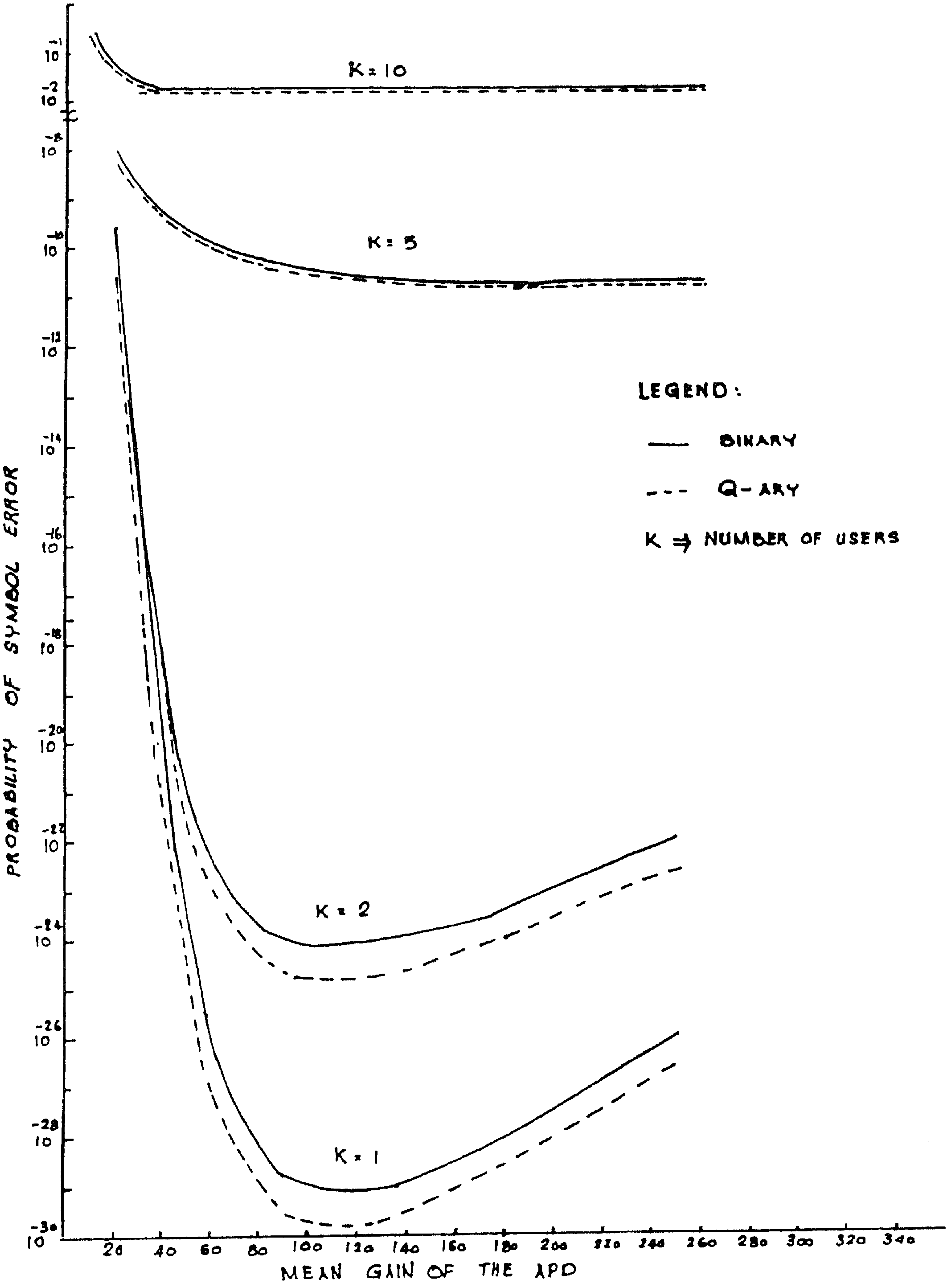


FIG3. EFFECT OF MEAN GAIN OF APD ON THE PROBABILITY OF SYMBOL ERROR. N = 341, r = 5

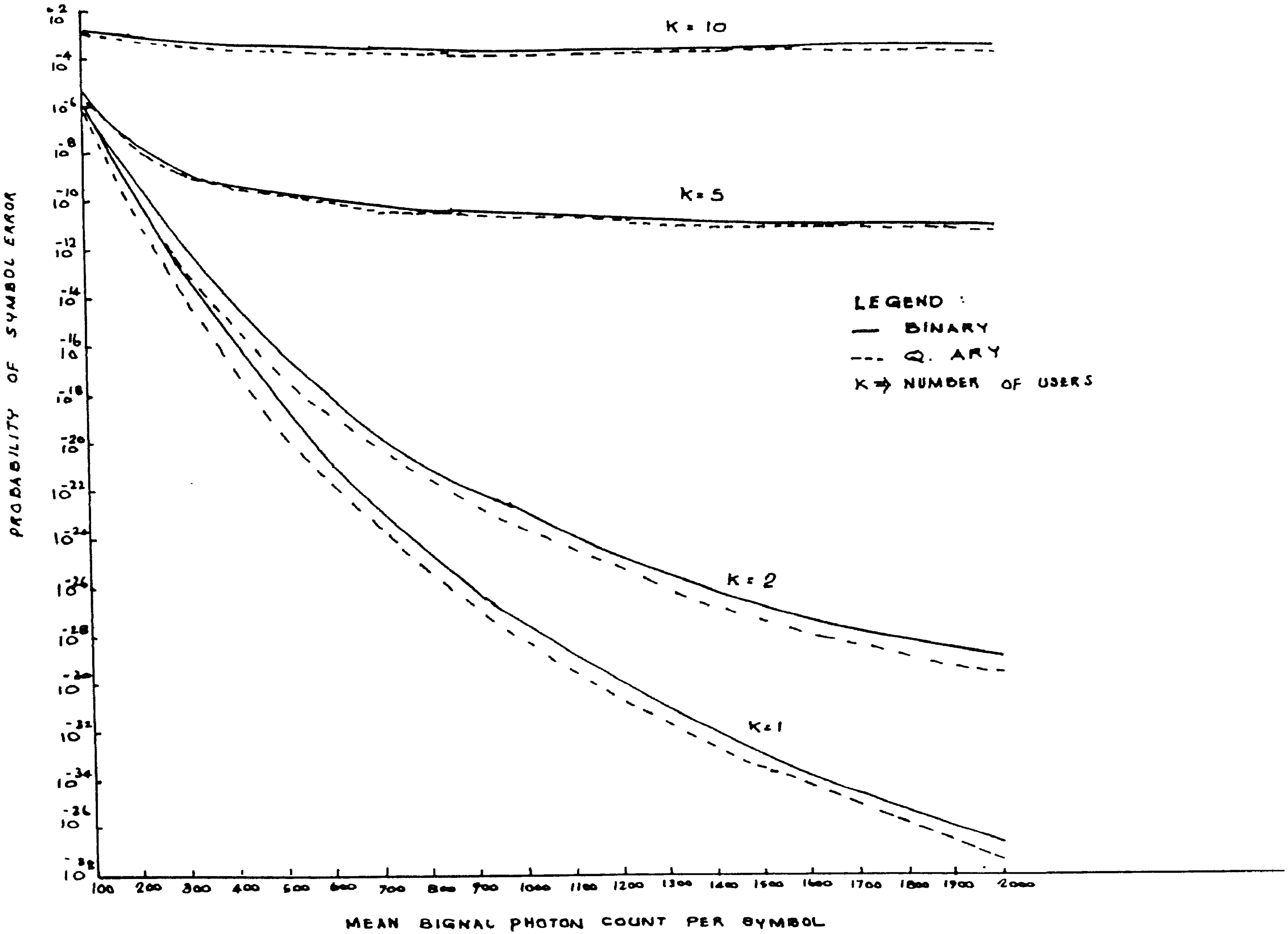


FIG. 4. EFFECT OF THE MEAN SIGNAL PHOTON COUNT PER SYMBOL ON THE PROBABILITY OF SYMBOL ERROR. N = 341, r = 5

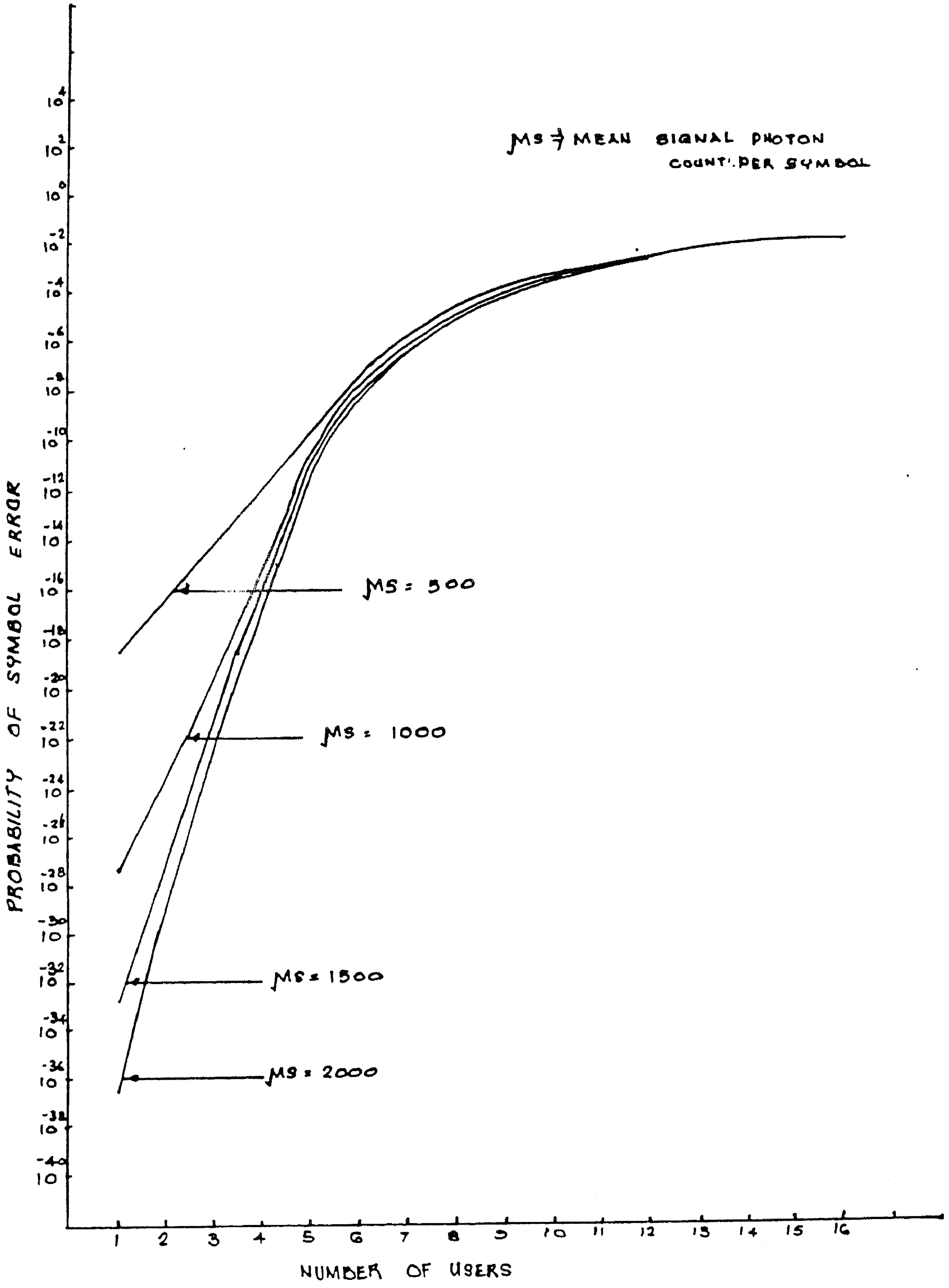


FIG.5 EFFECT OF THE NUMBER OF USERS ON THE PROBABILITY OF SYMBOL ERROR : N = 341, r = 5

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 135-146

FUZZY ENTROPY NET : APPLICATION TO RECOGNITION OF HANDPRINTED ORIYA CHARACTERS

R.Panda and S.Patnaik
Dept. of Electronics and Telecommunication Engg.,
University College of Engg., BURLA, INDIA - 768018.

G.S.Rath
Dept. of Applied Electronics and Instrumentation Engg.,
Regional Engineering College, Rourkela, INDIA - 769008.

Abstract- This paper presents the design of a class of fuzzy neural network, called fuzzy entropy net and its application to recognition of handprinted Oriya characters. A new methodology for network training employing fuzzy entropy is introduced. Spline transform of the digitized X- and Y- signature waveforms of a character are taken in evaluating the ellipse parameters, which are used as shape descriptors. Elliptic properties of spline coefficients are shown and expressions for the major axis and the minor axis of the k-th component ellipse are derived. The fuzzy membership function is defined and used for calculation of fuzzy entropy.

I. INTRODUCTION

The fascination of todays multilayered artificial neural nets (ANN) is that it is capable of classifying different classes in a multifeature multidimensional space [1]. I.K.Sethi [2] has presented a two step methodology for designing entropy networks, i.e., the mapping of decision trees into a multilayer neural network. He has claimed that the above methodology is superior than the traditional classifiers in the sense that it specifies a fixed number of neurons needed in each layer along with the desired input and it leads to a faster progressive training procedure that shows each layer to be trained separately.

In this paper, we have proposed a fuzzy entropy net and its application to recognition of handprinted Oriya characters. A new methodology for network training employing fuzzy entropy is introduced. Fuzzy entropy is deeply associated with information and conditioning (learning) and is no way related to probability theory [3-5]. We have adopted the fuzzy entropy measure proposed by De Luca and Termini [4] , which is quite intuitive and stuck in the fuzzy literature. The spline transform proved to be a powerful tool for approximating any function within an interval in the sense of minimum mean square error [6.7]. Hence, it is used for evaluating the spline descriptors of Oriya characters. The Fourier descriptors obtained by Persoon and Fu's [8] method of double projection profile requires less number of frequency components while features are used in pair. Tai et al [9] and Kuhl et al [10] have suggested a set of component ellipses as shape descriptors, where features were also used in pair. A similar method of using ellipse parameters as shape descriptors for the recognition of Oriya characters is adopted. But the main difference of the method suggested is that the spline transform of the double projection profile is used instead of Fourier transform. The fuzzy membership function (in terms of the major axis and the minor axis of both the standard and handprinted characters) is defined and used for calculation of fuzzy entropy of a character.

II. FUZZY ENTROPY NET

It is clear from the literature that there exist several tree generation procedures that are driven by the example pattern vector [2]. Any one of these procedures can be used to design a fuzzy entropy network for a given pattern recognition problem. The design of such network provides a direct method of classification of a multifeature multidimensional space. In a decision tree, if two or more leaf nodes result in the same decision, then the corresponding paths are in OR relationship. A layered artificial neural network implements ANDing of hyperplanes followed by ORing in the output layer. Therefore, one can comment that a layered ANN and a decision tree are equivalent. AND/OR relation in binary logic is equivalent to MIN/MAX relation in fuzzy logic. An example of a fuzzy entropy net for classification of three unknown classes is shown in Fig.1. There are three layers called as the input layer, MIN layer and MAX layer. The mapped fuzzy entropy net shown in Fig.1 has far fewer connections compared to the standard feedforward layer network. The mapped network described in this section is called as fuzzy entropy net due to its classification capability based on fuzzy entropy. The present mapping rule transform a sequential decision making process into a highly parallel process. Output from a node of the MIN layer is given by

$$n_{iout} = \begin{cases} x & \text{if MIN } (x, h_i+\delta) = x \\ h_i+\delta & \text{otherwise} \end{cases}$$

where x is the fuzzy entropy of unknown pattern class

h_i is the measured entropy of known pattern class

and δ is the threshold value

The number of nodes in both the input and output layer equals the number of distinct classes. Output from a node of the MAX layer is

$$n_{iout} = \begin{cases} \text{input 1} & \text{if MAX (input 1, input 2)=input 1} \\ \text{input 2} & \text{otherwise} \end{cases}$$

where input 1 and input 2 are the outputs from a node of MIN layer and a node of input layer, respectively.

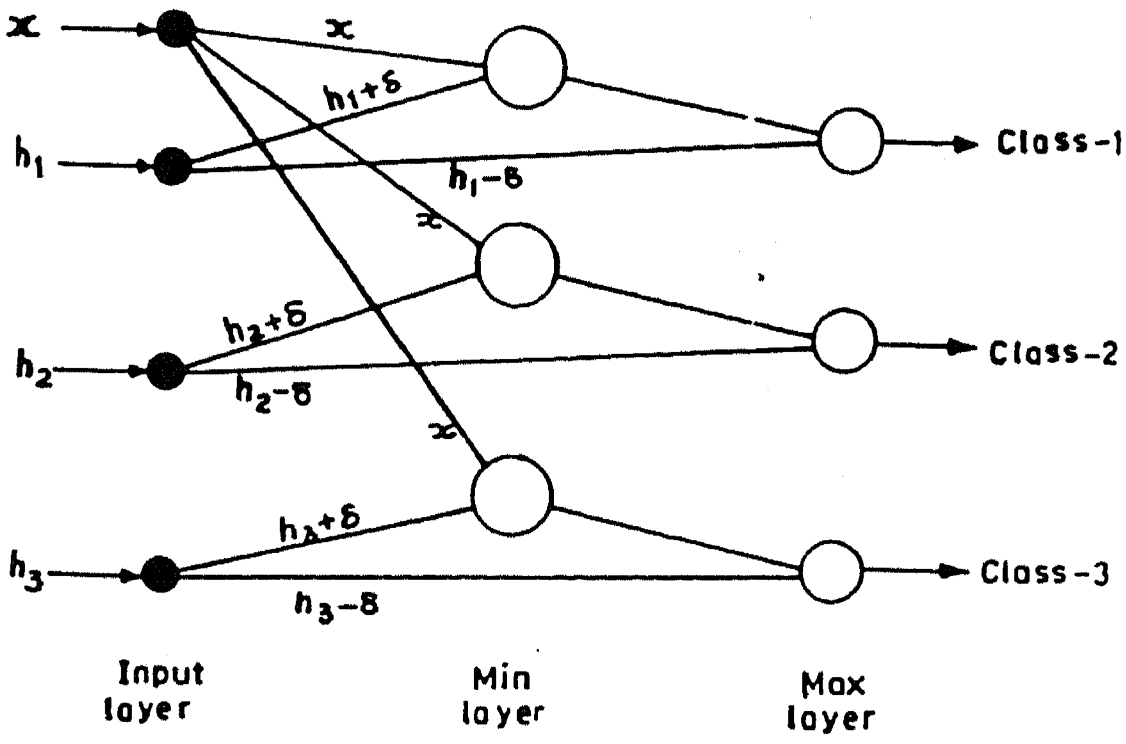


Fig.1. Fuzzy Entropy Net

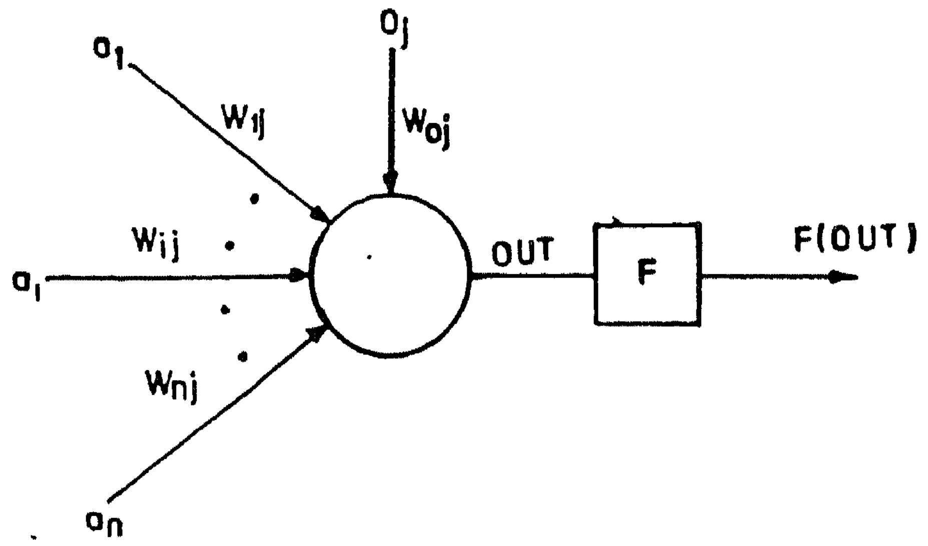
When the output from a node of MAX layer is x, then the decision is taken and the node is declared as a terminal node; otherwise it is considered as a nonterminal node. Fig.2 displays the anatomy of a generic processing element (PE), which can be considered as a node in the above fuzzy entropy net. The output computation is performed by taking the weighted sum, subtracting the threshold, and passing the result through a threshold function f( ). Mathematically this operation is defined as

$$OUT_j = f\left(\sum_{i=1}^{n} a_i \omega_{ij} - \omega_{0j} \theta_j \right) \qquad (1)$$

The nonlinear ramp threshold function shown in Fig.3 is used to get unit interval output values. The threshold function is described by the equation

$$f(x) = \begin{cases} +\gamma & \text{if } x > \gamma \\ x & \text{if } |x| < \gamma \\ -\gamma & \text{if } x < \gamma \end{cases} \qquad (2)$$

where $\gamma(-\gamma)$ is the PE's maximum (minimum) output value. Although (2) is a piece-wise linear function, it is often used to represent a simplified nonlinear operation. Each layer of the above fuzzy entropy net has been trained separately.



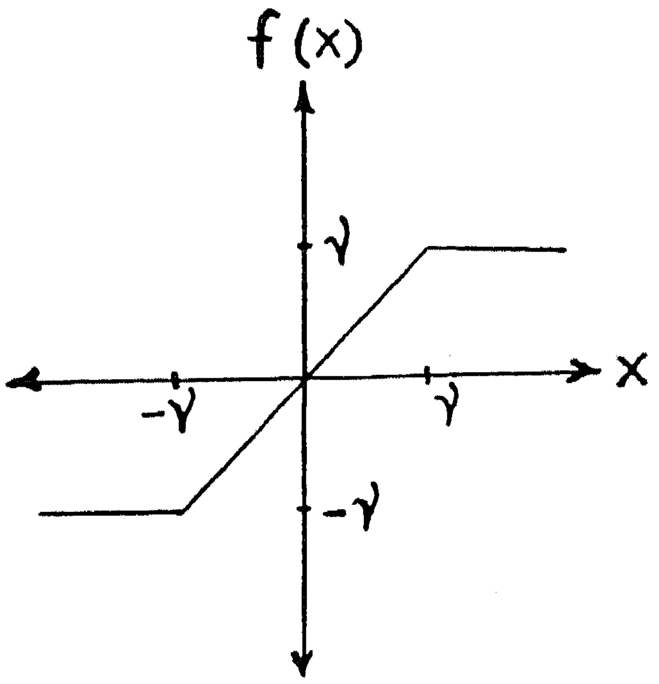


Fig.2. The Basic PE

Fig.3. Threshold Function

III. SPLINE INTERPOLATION AND SPLINE TRANSFORM

In an interpolation scheme, the interpolated function is given as

$$\hat{f}(t) = \sum_{k=1}^{k} C_k S_k(t) \qquad (3)$$

where C_k are the coefficients to be determined from the input data, $S_k(t)$ are basis functions and k is the number of given data points.

When $S_k(t)$; k=1,2,3,......,k; a subset of set of orthogonal functions ; are the spline functions; then the scheme is known as spline interpolation scheme. The principle of spline interpolation of any function by B-spline functions has been utilized to develop spline transform [6,7]. R.Panda and G.S.Rath [7] have presented the theoretical investigation of impulse response identification using B-spline series functions and discrete spline transform. The discrete spline transform of a data sequence X(m), m=0,1,2,......,N-1 is defined as [6,7]

$$S(2k-1) = \frac{T}{N} \; \frac{n!}{\lambda_m^{(n)}} \sum_{m=0}^{N-1} X(m)\, Y_{2k-1}^{(m)} \qquad \text{for } k=M \text{ when } N \text{ is even}$$

$$= \frac{2T}{N} \; \frac{n!}{\lambda_k^{(n)}} \sum_{m=0}^{N-1} X(m).Y_{2k-1}^{(m)} \qquad \begin{array}{l} \text{for } k=1,2,\ldots\ M-1 \\ \text{when } N \text{ is even} \\ \text{or } k=1,2,\ldots.\ M \\ \text{when } N \text{ is odd} \end{array} \qquad (4)$$

and

$$S(2k) = \frac{T}{N} \sum_{m=0}^{N-1} X(m) \qquad \text{for } k=0$$

$$= \frac{2T}{N} \; \frac{n!}{\lambda_k^{(n)}} \sum_{m=0}^{N-1} X(m).Y_{2k}^{(m)} \qquad \begin{array}{l} \text{for } k=1,2,\ldots\ M-1 \\ \text{when } N \text{ is even} \\ \text{or } k=1,2,\ldots.\ M \\ \text{when } N \text{ is odd} \end{array} \qquad (5)$$

where S(k) is the k-th spline transform coefficient, T is the sampling interval, n is the degree of B-spline, $Y_k^{(i)}$ is the i-th element of the k-th eigen vector$[Y_k]$, $\lambda_k^{(n)}$ is the k-th eigen value of the system matrix $[A^{(n)}]$ for the periodic B-spline function, and M is the integer division of N by 2.

IV. SPLINE DESCRIPTORS OF ORIYA CHARACTERS

In shape analysis, the relevant information about the shape of the object is extracted in compact form from the binary or gray level image of the object. The features that describe the shape in a compact form are known as descriptors. The shape descriptors should exhibit good discrimination capability. In this section, a method of shape discrimination using spline descriptors is presented. In this technique, the boundary of the character is converted into a pair of periodic waveforms and consists of projection on the X-axis and Y-axis, respectively. Each projection profile designates the change in the coordinate (X-or Y-) value as a function of arc length t measured from a chosen starting point in the contour. It is a smooth function of t without any sharp discontinuity.

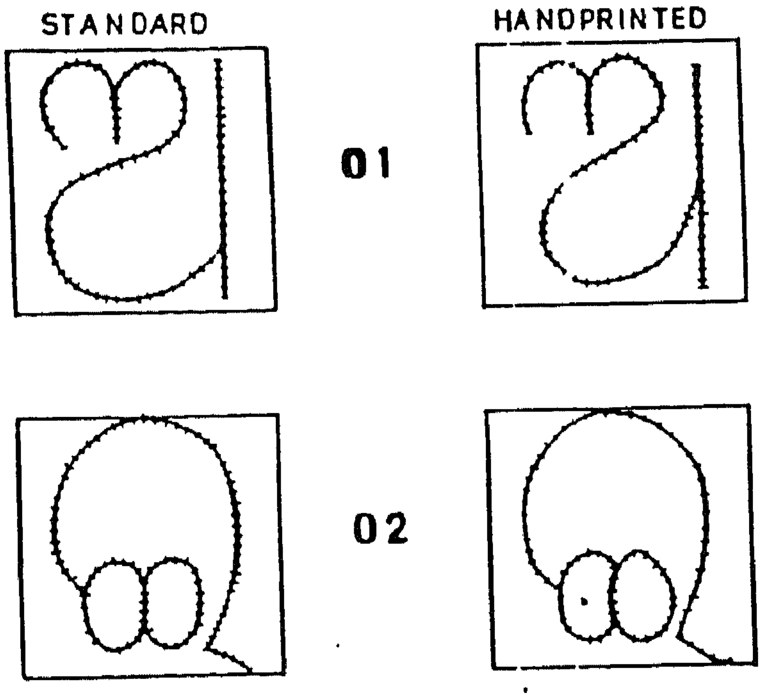


Fig.4.Standard & Handprinted characters '01 & 02' used for digitization

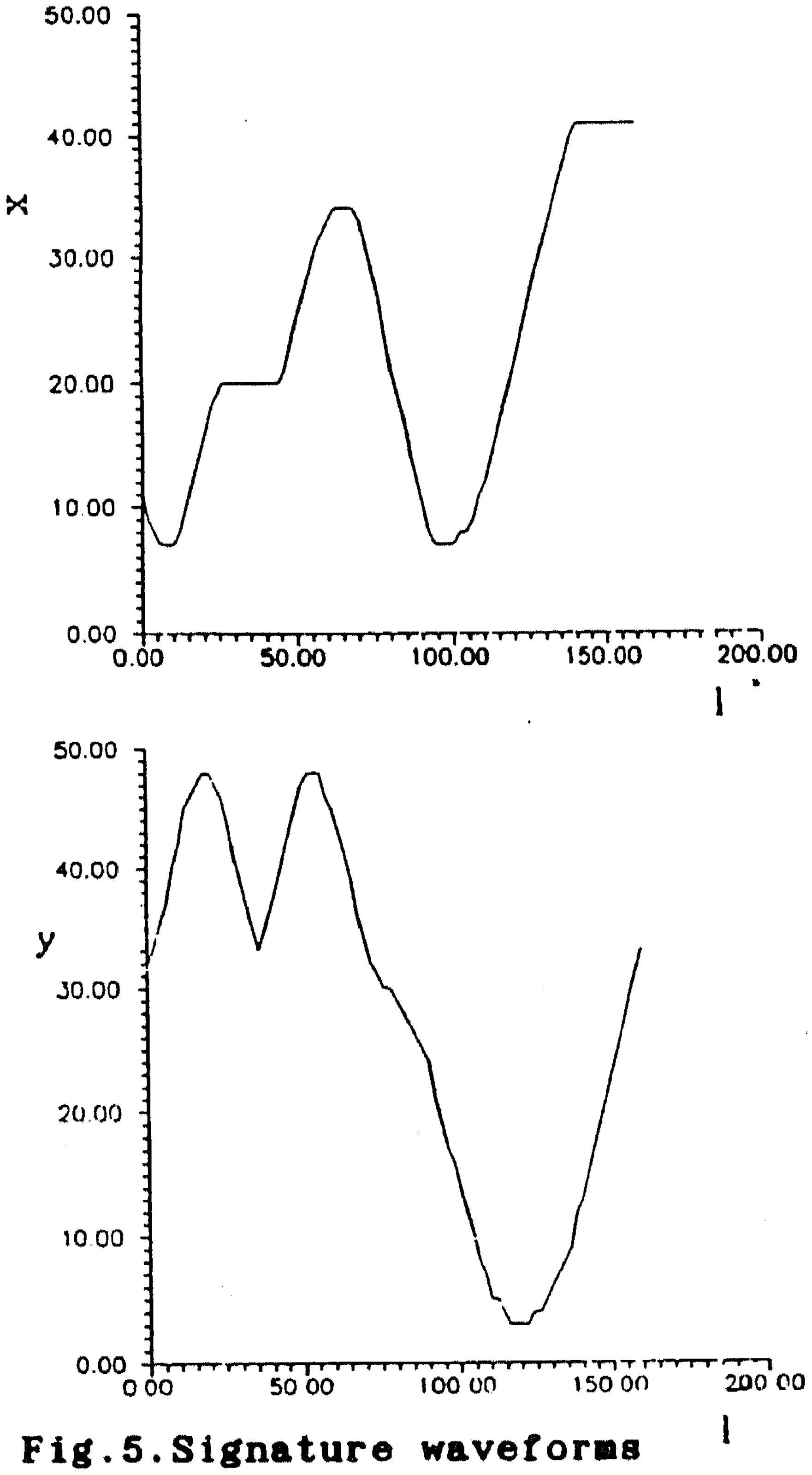


Fig.5.Signature waveforms for 01-character (standard)

Fig.4 illustrates the method of tracing and digitizing the x- and y- coordinates of the boundary of characters. In this figure, the actual digitization of both the standard and handprinted characters (01 and 02) has been shown. The digitized signature waveforms of 01-character are shown in Fig.5.

The digitized X- and Y- signature waveforms are taken in evaluating the ellipse parameters.Using the generalized truncated fourier series, the X- and Y- signature waveform can be represented by

$$X_k \cong \sigma_k(0)+ \sum_{i=1}^{L} \sigma_x(2i-1)\mathrm{Sin}(2k-1)i\theta+\sum_{i=1}^{L} \sigma_x(2i)\mathrm{Cos}(2k-1)i\theta \qquad (6)$$

and

$$Y_k \cong \sigma_y(0)+ \sum_{i=1}^{L} \sigma_y(2i-1)\mathrm{Sin}(2k-1)i\theta+\sum_{i=1}^{L} \sigma_y(2i)\mathrm{Cos}(2k-1)i\theta \qquad (7)$$

where L indicates the highest harmonic component, $\sigma_x(k)$ is the k-th transformed coefficient of x-signature waveform and $\sigma_y(k)$ is the k-th transformed coefficient of y-signature waveform and $\theta = \pi/N$. when N is large (usually N is greater than 90) X_k and Y_k approximate to the continuous functions of θ. The average components $\sigma_x(0)$ and $\sigma_y(0)$ represent the location of the geometric centre of the contour which do not contain any information of the shape. So they are ignored while evaluating ellipse parameters. Let X(k) and Y(k) be the coordinates of the curve represented by k-th spline harmonic component and are given as

$$X(k) = \sigma_x(2k-1)\ \mathrm{Sin}\ k\alpha + \sigma_x(2k)\ \mathrm{Cos}\ k\alpha \qquad (8)$$

and $$Y(k) = \sigma_y(2k-1)\ \mathrm{Sin}\ k\alpha + \sigma_y(2k)\ \mathrm{Cos}\ k\alpha \qquad (9)$$

where α denotes the continuous, angular function corresponding to the discrete version of θ and $0 \le \alpha \le 2\pi$. (8) & (9) represents an ellipse in parametric form. (8) & (9) when transformed to polar form with radius vector

$R_k(R^2=\sqrt{x^2+y^2})$ and polar angle ψ_k, becomes,

$$R_k^2 = P_k + \sqrt{(P_k^2 - Q_k^2)} \sin (2k\alpha+\nu_k) \tag{10}$$

where

$$P_k^2 = \frac{\sigma_x^2(2k-1)+\sigma_x^2(2k)+\sigma_y^2(2k-1)+\sigma_y^2(2k)}{2}$$

$$Q_k = \sigma_x(2k)\sigma_y(2k-1)-\sigma_x(2k-1)\sigma_y(2k)$$

$$\nu_k = \mathrm{Tan}^{-1} \frac{P_k}{\sigma_x(2k-1)\ \sigma_x(2k)\ +\ \sigma_y(2k-1)\ \sigma_y(2k)}$$

and

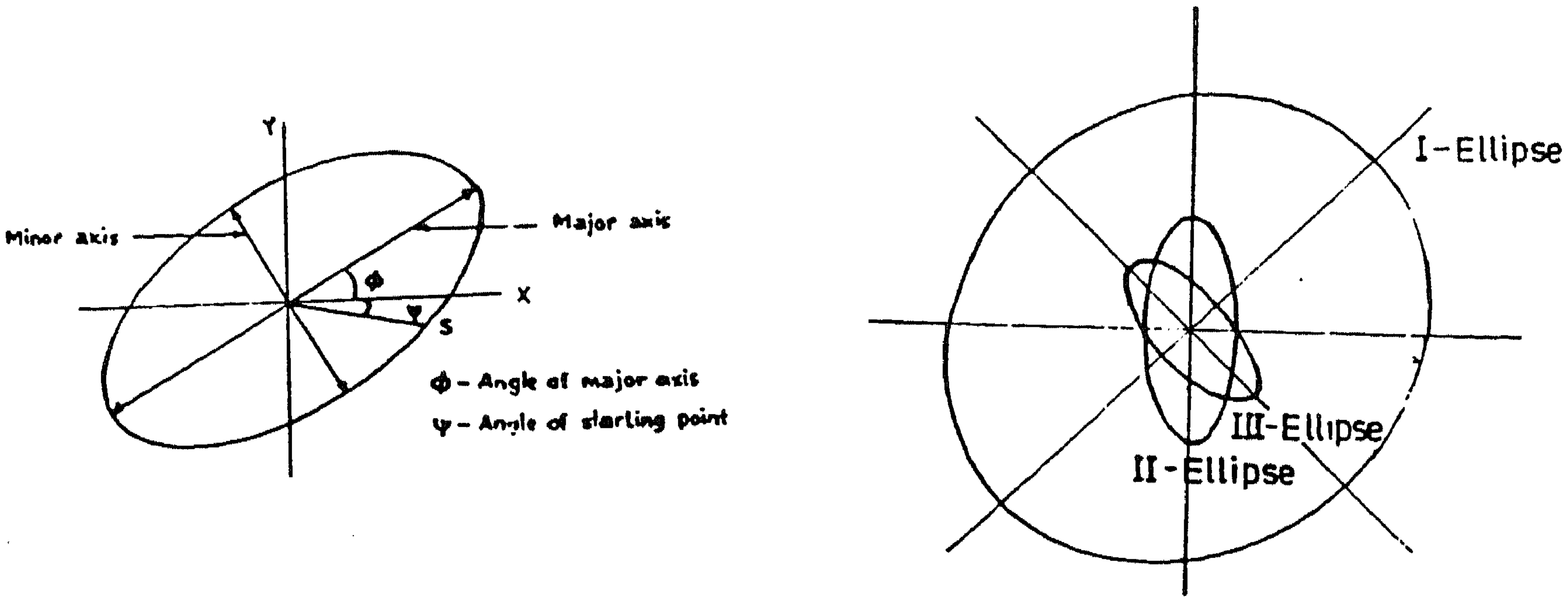
$$\psi_k = \mathrm{Tan}^{-1} \left. \frac{Y(k)}{X(k)} \right|_{t=0}$$

Using (10), the major & minor axis of the ellipse can be obtained as

$$\delta_{maj}(k) = 2\ [P_k + \sqrt{P_k^2 - Q_k^2}\ ]^{1/2} \tag{11}$$

$$\delta_{min}(k) = 2\ [P_k - \sqrt{P_k^2 - Q_k^2}\ ]^{1/2} \tag{12}$$

where $\delta_{maj}(k)$ and $\delta_{min}(k)$ are the major and the minor axis of the k-th component ellipse. The parametric angle of the major axis is $\frac{\pi/2-\nu_k}{2k}$. Therefore, the k-th component ellipse is completely described by the major axis, the minor axis, parametric angle of the major axis and the polar angle of the starting point. Although the regeneration of an approximation of original contour requires the information of all four parameters of a set of component ellipses, the major axis and the minor axis of the component ellipse are taken as shape descriptors. These shape descriptors provide an intuitively appealing interpretation acceptable to human perception of shapes. The parameters of a component ellipse and the first three component ellipses of O1-character are shown in Fig.6.



(a) (b)

Fig.6. (a) The component ellipse for the k-th spline harmonic of signature waveforms,(b) First three component ellipses for O1-char.

V. FUZZY ENTROPY MEASURE BASED ON ELLIPTIC SPLINE FEATURES

The ellipse parameters, i.e., the major axis and the minor axis in pair, are taken as shape descriptors for classification. The fuzzy membership function, used for measuring fuzzy entropy, is defined as

$$\mu_A(x_i) = \min \left[\frac{\delta_{smin}}{\delta_{smaj}} , \frac{\delta_{hmin}}{\delta_{hmaj}} \right] \tag{13}$$

Where δ_{smin} is the minor axis of the standard character
δ_{smaj} is the major axis of the standard character
δ_{hmin} is the minor axis of the handprinted character
δ_{jmaj} is the major axis of the handprinted character

The fuzzy entropy of a handprinted character is defined as [4]

$$H(A) = (1/n \ln 2) \sum_i - \mu_A(x_i) \ln \mu_A(x_i) - \{1-\mu_A(x_i)\} \ln \{1-\mu_A(x_i)\} \tag{14}$$

where $\mu_A(x_i)$ is the fuzzy membership function defined in (13).

VI. CLASSIFICATION OF ORIYA CHARACTERS

The Oriya (appropriate pronunciation is Odiya) is one of the fifteen major Indian languages recognized by the Constitution of India, used by more than forty five million people. Oriya is an offspring of Sanskrit and its grammar has close resemblance with that of Sanskrit. Its alphabet consists of forty nine simple characters and many compound characters. Out of forty nine simple characters, there are thirteen vowels and thirty six consonants. A group of characters in isolation forms a word similar to Greek and German languages. There is no concept of upper case and lower case and the script is written in a uniform size. The Oriya characters are predominantly round in shape. Thirty five (five vowels and thirty consonants) mostly used characters are considered for recognition problem. All thirty five characters with their character codes are shown in Fig.7. In this classification experiment we have considered first ten component ellipses only. Fuzzy entropy values of all thirty five characters have been calculated using (14) and are presented in Table-1. The value of threshold δ is chosen to be 0.002. A complete fuzzy entropy net has been simulated and used for classification of all thirty five handprinted Oriya characters. We have taken 100 samples for the present recognition problem. The percentage of correct classification is found to be 93%.

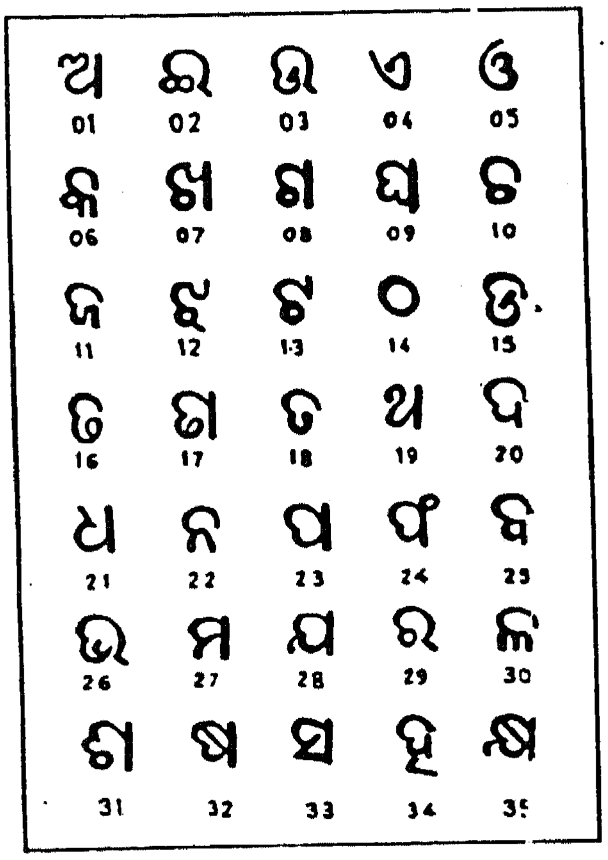


Fig.7. 35 mostly used Oriya characters with code

TABLE 1 : FUZZY ENTROPY VALUES

CHAR. CODE	FUZZY ENTROPY VALUES	CHAR. CODE	FUZZY ENTROPY VALUES	CHAR. CODE	FUZZY ENTROPY VALUES
01	0.7738	13	0.5835	25	0.5665
02	0.4597	14	0.3369	26	0.6480
03	0.7284	15	0.5151	27	0.6053
04	0.5022	16	0.6890	28	0.6529
05	0.4944	17	0.6567	29	0.6557
06	0.6001	18	0.6248	30	0.5704
07	0.7757	19	0.6089	31	0.7657
08	0.6660	20	0.5886	32	0.7213
09	0.7073	21	0.6636	33	0.6769
10	0.5730	22	0.6599	34	0.6454
11	0.5762	23	0.5334	35	0.7530
12	0.6155	24	0.7326		

VII. CONCLUSION

Since an overview of shape analysis stresses mainly on the boundary based global technique, the same is used for the present recognition problem. The shape descriptors represented by the major axis and the minor axis provide intuitively more faithful information regarding the shape. Such descriptors are quite useful for Oriya characters because they have more circular (rounded) appearance. This is also confirmed from the recognition results. The spline transform is a powerful tool as it does not involve complex multiplication like DFT; it is quite similar to discrete cosine transform (DCT) and discrete sine transform (DST). Moreover, the signal can be reconstructed more easily through the spline transform. Hence, it is used to extract the spline features. The proposed fuzzy entropy net has far fewer connections compared to the standard feedforward artificial neural nets, which is an attractive feature for VLSI implementation point of view. The present recognition problem may be extended to complete set of characters including compound characters.

REFERENCES

[1] R.P.Lippmann , "An Introduction to computing with Neural Nets," IEEE Trans. ASSP Mag. , pp. 4-22 , April 1987.

[2] I.K.Sethi, "Entropy Nets : From Decision Trees to Neural Networks," IEEE proc. , vol. 78, pp. 1605-1613, October 1990.

[3] A.Kaufmann , "Introduction to the theory of fuzzy subsets," vol. 1, Academic Press, New York, 1975.

[4] A.De.Luca and S.Termini , "A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory," Information and Control , vol. 20 , pp. 301-312 , 1972.

[5] Bart Kosko , "Fuzzy Entropy and Conditioning ," Information Science, vol. 40 , pp. 165-174, 1986.

[6] G.S.Rath, "On splines, spline transform and application to handprinted Oyiya character recognition," Ph.D Thesis , I.I.T. Kharagpur, India , Nov. 1984.

[7] R.Panda and G.S.Rath, "Computer oriented approach to Impulse Response Identification ," Journal of IE (India) , ET, vol. 73, pp. 70-73 ,August 1992.

[8] E.Persoon and K.S.Fu , "Shape discrimination using Fourier descriptors," IEEE Trans. Systems, Man, Cybern. , vol. SMC-7 , pp. 170-179, Mar. 1977.

[9] H.T.Tai , C.C.Li and S.H.Chiang , "Application of Fourier descriptors to classification of fine particles ," Proc. 6-th International Conference on Pattern Recognition , Munich, pp. 748-751 , 1982.

[10] F.P.Kuhl and C.R.Giardina , "Elliptic Fourier features of a closed contour ," Computer Graphics Image Processing , vol.18, pp. 236-258 , 1982.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 147-154

MODIFIED WINDOW FUNCTION FOR FINITE-IMPULSE-RESPONSE DIGITAL FILTERS AND DATA DETECTION

Majed A. Al Rasheedi[*]
Abdel-Naser Mansour
A. Raouf mohamed
Mohamed Galal Ali

Abstract

In this paper we shall study Raouf's and Nasser's window characteristics based on the graphical optimization technique. The study of the window function curve at different window orders and different values of its exponential parameter reported zero values at the window last terms. We shall introduce a modified window function that proved to be more efficient when it is used in conjunction with the Fourier series for digital filter design and implementation. The modified window parameters and their effects on the filter characteristics have been studied. The proper choice of the window parameters (optimal values) produced very promising FIR digital filter characteristics.

Introduction

Signals arise in almost every field of science and engineering ,e.g., in caustics, biomedical, communications, geophysics, solid state ,and nuclear physics. Two general classes of signals can be identified, namely continuous-time and discrete-time signals.

A continuous-time signal is one that is defined at every instant of time. A discrete-time signal is one that is defined at discrete instants of time such as microseconds, milliseconds and days. In fact the discrete-time signal, as if a continuous-time signal can be represented by a unique function of frequency referred to as the spectrum of the signal . Practically, the transformation processes from any continuous-time system into a discrete-time system introduce two major sources of errors. The first source comes from the quantization of the system data and system transfer function, and the second source introduced from applying Fourier transformation to the system data with a finite number of coefficients for practical implementation [1,2,3,4].

---

[*]Demonstrator ,Applied Science Dept. , College of Technological Studies , Kuwait .

Also the representation of Fourier series coefficients in discrete form will contribute significant errors when the digital system word-length is not sufficient. However, the quantization effect on the digital processing is not of our interest now. Many groups concentrated their works on the problem arises from the limitation of the number of Fourier coefficients in the FIR digital filter transfer function and its characteristics [2,5,6,7]. Window technique is one of the most important and powerful techniques selected to compensate of the limitation number of the Fourier coefficients .A Considerable numbers of windows have been presented during the last 15 years. Those windows are constructed based on many different optimization constraints.

Raouf and Naser used technique named Graphical Optimization Technique to construct their window that proved to be sufficient in the FIR digital filter design and implementation. In this paper we shall study Raouf's window in more details to explore the real correlation between the filter order N, and the window function parameter "a" moreover their effect on the FIR filters characteristic curve. Also we shall introduce a modified window to get better filter response and to avoid the problem raised in Naser window.

Theoretical Approach to modify Raouf's Window

Raouf constructed his window function which has three parameters. This window is given as:

$$W(n)=\mathrm{Exp}(-a\,\pi\frac{n^2}{N})\,\mathrm{Cos}(b\,\pi\frac{n}{N})\,(\frac{1}{1+(c\,n^4)}) \qquad (1)$$

Where n=0,1,2,...,N-1 and a, b, c are constant parameters. These parameters have been studied based on the graphical optimization technique. The study proved that the window function is not sensitive to parameters "b" and "c" .The optimized values of parameters "b" and "c" where found to be 0.47 and $(4*10^{-5})$ respectively [8]. Then, the window shape is now entirely dependent on parameter "a" and the window order N ,

$$N=(\text{Filter order-1})/2 \qquad (2)$$

Raouf and Naser studied the dependence of the FIR filter response on parameter "a" for different filter orders. They also found that they can have the desired FIR digital filter characteristic at different orders by changing the value of parameter "a". This means that the required filter could be obtained at low order if the parameter "a" chosen properly.

Figure (1) shows Raouf's window shape at different orders with different values of parameter "a". It can be seen that, as the window function order progressively increases, the outlying function values rapidly approach zero. Therefore, when we do a filter implementation at specific high order value, we really got a filter response corresponding to lower filter order, i.e., at high window function orders, there is a difference in the observed vs. expected which is inevitable value due to the phenomenon resulting from the zero values and narrow width of the window function.

To understand this result we shall have to go back to equation (1). In this equation the exponential term has the ratio (a/N) which remains constant ratio to give the same response as the values of "a" and "N" change in fixed proportion. At higher order of the FIR digital filter, the last term in equation (1) strongly affects the values of the window function and makes its last values equal zero. Therefore, it is quite logic to think about omitting the window last term and try to compensate for it through the first window term (exponential term).

Thus relying on the Graphical optimization technique and considering the last window values equal zeros ,as the order of the filter increases, and aiming to simplify the structure of Raouf's window, we propose the following modified window function

$$W(n)=\mathrm{Exp}(-A\,\pi\,\frac{n^2}{N^2})\,\mathrm{Cos}(B\,\pi\,\frac{n}{N}) \qquad -N/2 \leq n \leq N/2 \qquad (3)$$

Where "A" and "B" are constant parameters .To select the two parameters "A" and "B" we shall use the Graphical Optimization technique and Nasar's results. It was found that the optimal value of "B" is given by

$$B=2b \quad , \; b=0.47 \qquad (4)$$

Study has been done to examine the effects of this new window function on the characteristic curves of the FIR filter response. The filter response was calculated at different values of parameter "A."The optimal value of "A" was found to be in the range given by :

$$A=2a \quad , \; 0.6 \leq a \leq 3 \qquad (5)$$

Figure (2) shows the filter response calculated at 4 different values of parameter "a". From this figure it can be seen that, as the value of "a" remains constant and the number of Fourier coefficients increased (filter order N), the height of the first side-lobe remains constant, whilst the main-lobe width decreases greatly. On the other hand, when the value of parameter "a" increase (constant N values), the height of the first side lobe decreases simultaneously with the decrease in main lobe width.

In order to compare the effect of this modified window function, and Raouf's window ,two different values of parameter "a" were chosen to calculate their FIR digital filter characteristics at order N equal 51 .Fig. (3) shows this comparison in which the parameter "a" for each window function was selected on by these values where their main-lobe width would be nearly equal. Clearly the modified window function has resulted a significant decrease in the height of the first side-lobe as compared to Raouf's window function.

Discussion and Conclusion

This Modified window function is a major breakthrough in providing responses with low height side-lobe and simultaneously narrow width main-lobe at low orders .Since these response characteristic's are independent of the order, but are rather directly related to the values of parameter "a" only. This feature of the modified window can be appropriately used to provide the desired characteristic of the FIR digital filter with minimum main-lobe width and very low height of the first side-lobe.

Also, This function has the distinct advantage of having a minimal ripple characteristic leading to optimal applicability to Ultrasonic implementations.

It is well known that Raouf and Kaiser's window nearly produce the same FIR digital filter characteristics at the same order, beside the simplicity of Raouf's window. Also from our results we have seen that the modified window has an advantage of providing more attenuation (approximately 35 dB) on the height of the first side lobe compared to Raouf's window, so the modified window has an advantage over the most well known windows.

Finally, based on our results we can strongly recommend this modified window for the implementation of digital filters and data detection as well as the Ultrasonic applications.

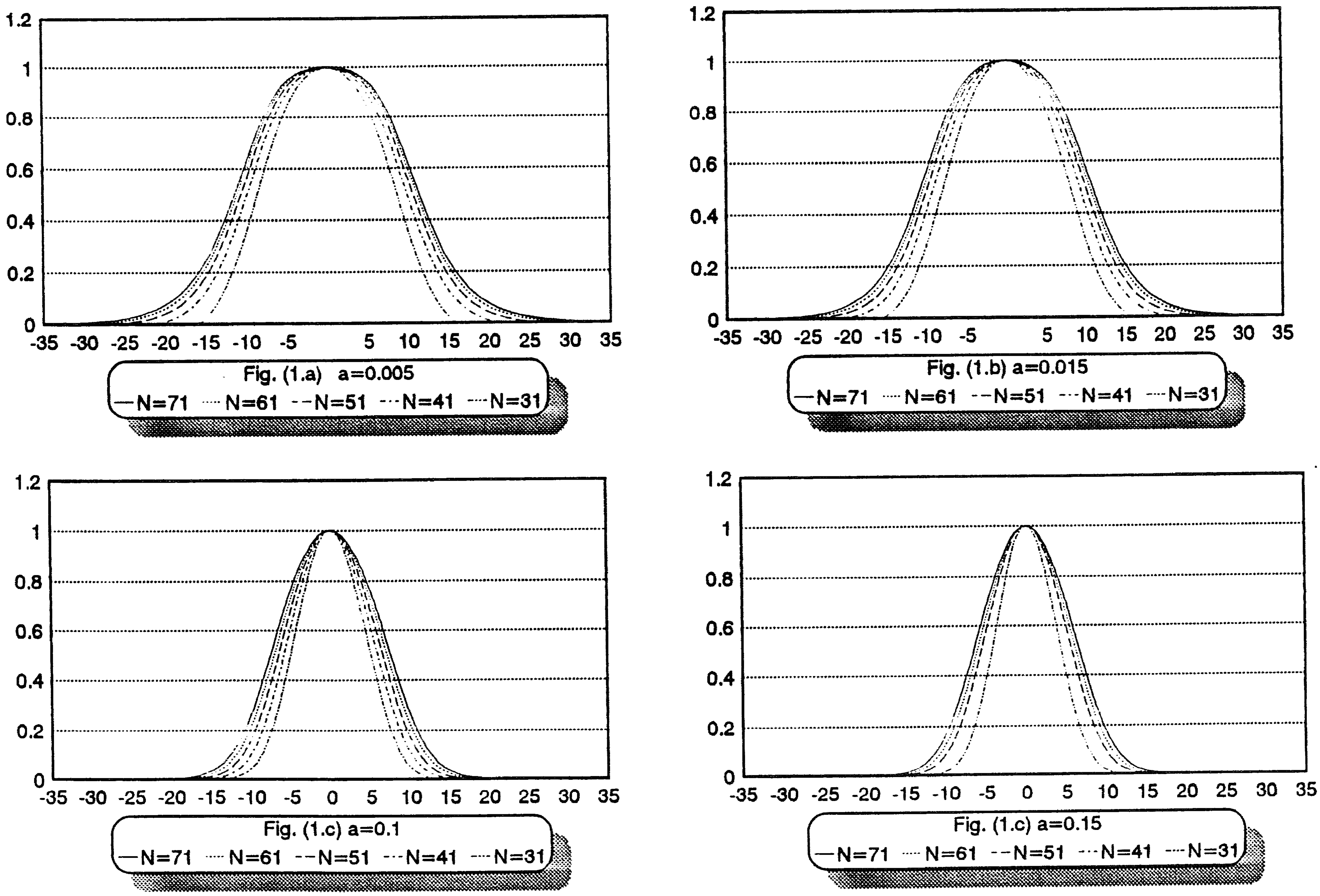


Fig. (1) Raouf's window with different window order for different values of parameter "a"

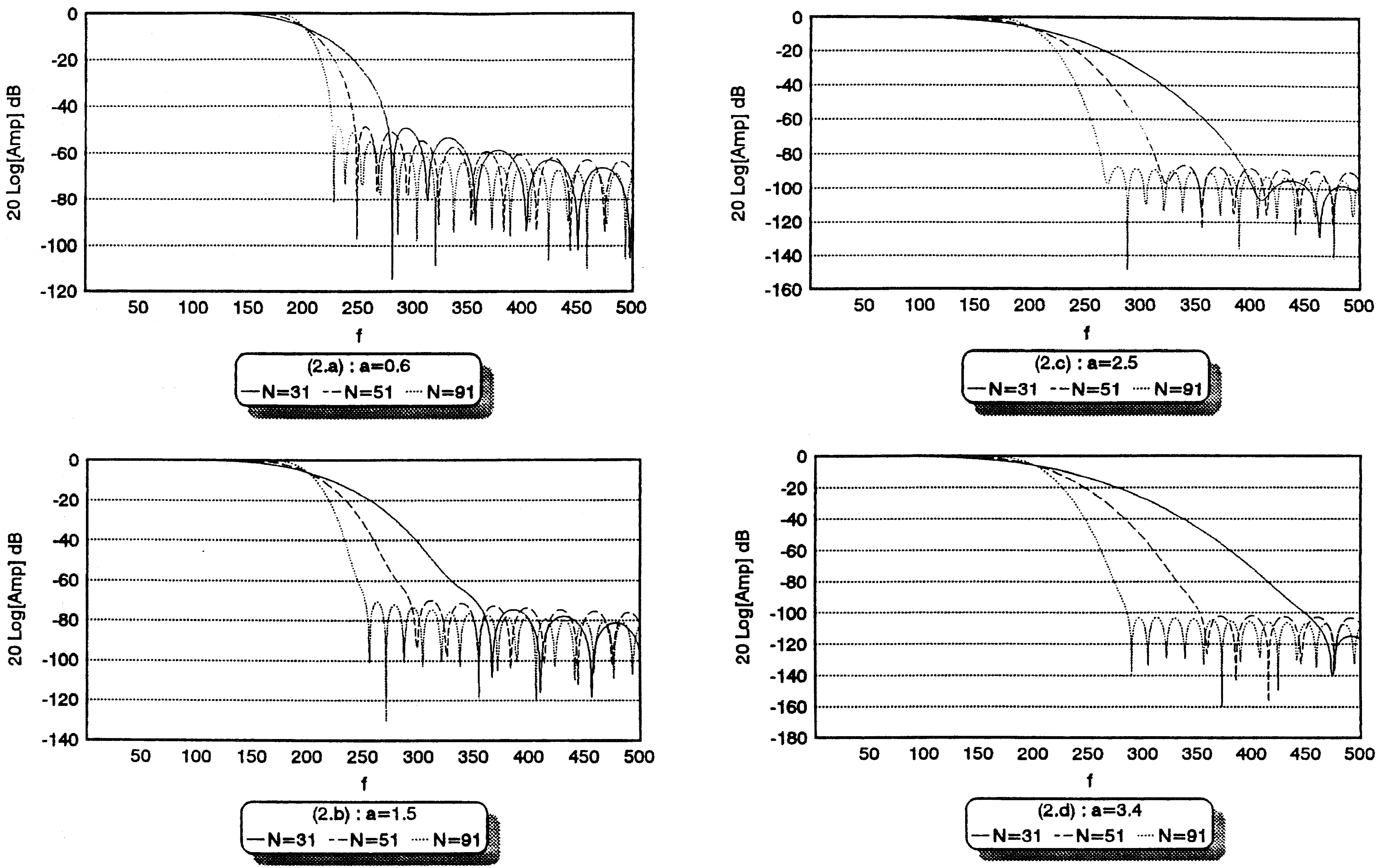


Fig.2 :FIR digital filter response for modified window at different values of parameter a and order N

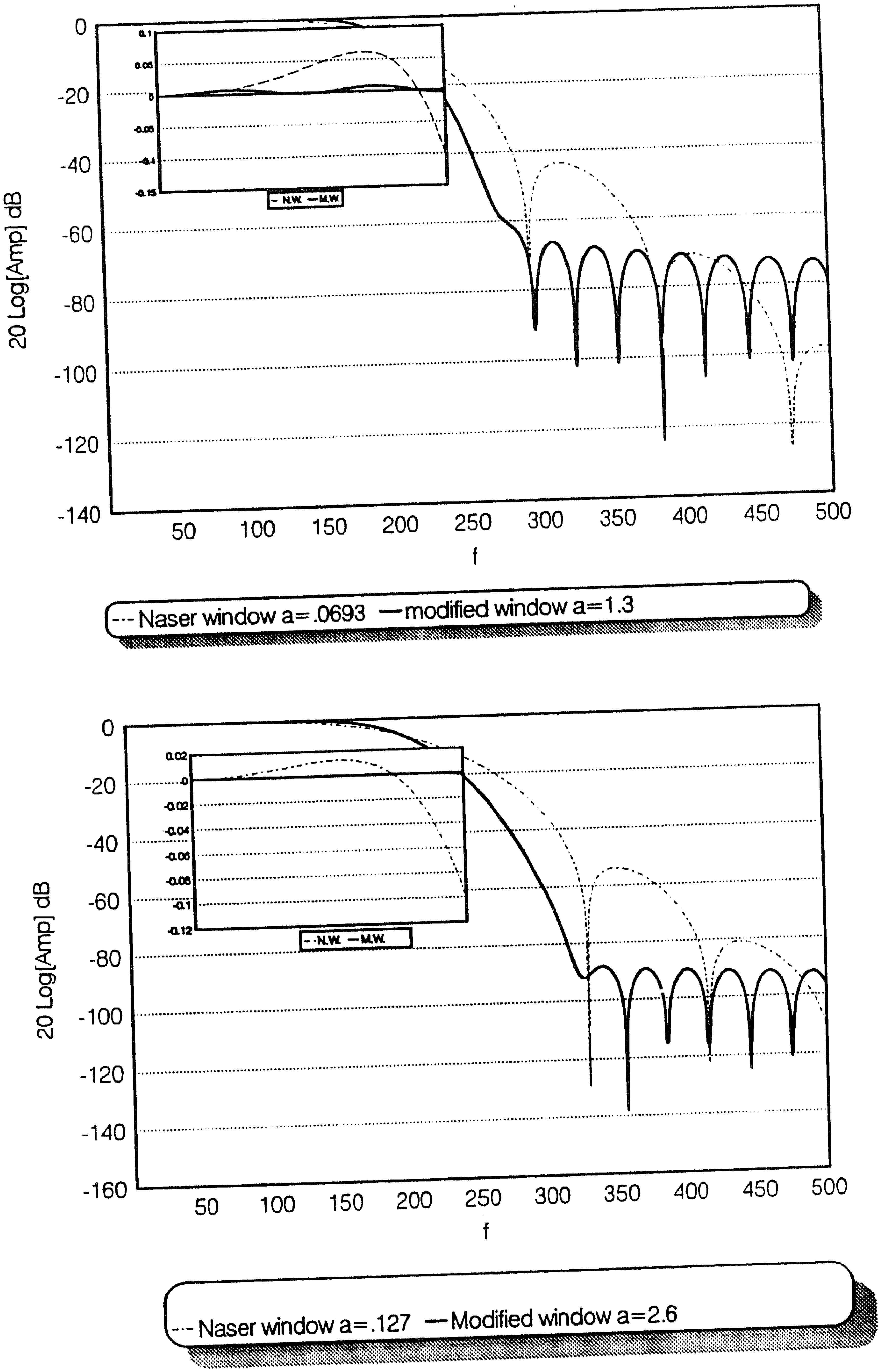


Fig.3 : Comparison between FIR filter responses for Naser and the modified windows at same main-lobe width.

REFERENCES

(1) R. E. Callis and R. I. Kitney: " The design of Digital filters for Biomedicl signals Processing, Part 1: Basic Concepts". J. Biomed. Eng. 1982 .Vol.5 .

(2) Ahmed Nasir: "Discrete-time Signal Processing". Prentice Hall, Englewood Cliffs, New Jersey. 1989 .

(3) R. W. Hamming:"Digital Filter". Prentice-Hall. Englewood Cliffs, New Jersey. 1989 .

(4) Robert W.Ramirez:"The FFT Fundemantals and Consepts".Prentice-Hall, Englewood Cliffs, New Jersey. 1980 .

(5) E. Oran Brigham: "The FFT and Its Applications".Prentice-Hall, Englewood Cliffs, New Jersey. 1988 .

(6) Nasser Abd El Azeem Kamel: "One and Two Dimensional FIR Digital Filter Design and Implementation".M.Sc. Thesis, Minia University, Cairo, A.R.E, 1989

(7) Esam Ahmad Mansour: "Graphical Optimization Technique in FIR Digital Filter". M.Sc. Thesis, Ain Shams University, Cairo, A.R.E.,1983.

(8) Nasir Ahmad: "Discrete-time Signals and System". Prentice-Hall, Englewood Cliffs, New Jersey .1989.

(9) Esam Ahmed Mansour: "Theoretical and experimental studies on the FIR filter",M.Sc. Thesis, Ain Shams University, Cairo, A.R.E, 1985 .

(10) A. Raouf Mohamed and N.A.Kamel: "Design and Implementaion of FIR Digital Filter Based On Window Parameters Optimization Technique". AMSE, Periodical. 1987.

(11) S. M. Bozic,"Digital and Kalman filtering".Edward Arnold.1979.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 155-165

A NEW APPROACH TO DECENTRALIZED FAULT DETECTION

K.A.GOPALA RAO
DEPT. OF ELECTRICAL ENGG.
ANDHRA UNIVERSITY
VISAKHAPATNAM

N.VISWANADHAM
DEPT. OF COMPUTER SCIENCE & AUTOMATION
INDIAN INSTITUTE OF SCIENCE
BANGALORE

ABSTRACT

In this paper, the decentralized fault detection (DFD) problem is formulated in frequency domain which will answer the following questions in a natural way:

1. How do we determine the minimum length of the window of observations that is required for detection and isolation?
2. Can we generate all possible parity vectors from a given minimum order parity vectors?

We illustrate the methodology with a numerical example.

1. INTRODUCTION

Failure detection and isolation (FDI) in dynamical systems deals with the problem of detecting deviations from normal behavior (we call them failure modes) in a specified complement of the system components (sensors, actuators etc.) and isolating the particular component that has failed. In recent years a wide variety of approaches to the problem of FDI in linear dynamical systems have been developed[1]. However, in these approaches the assumption is that all the information required for fault detection is available to the detector. This assumption does not hold in most practical situations, since for reasions of economy and reliability the information is spatially distributed. In many engineering systems the information available to fault detectors is thus localized and each detector monitors a specified section of the large system. This makes the Decentralized Fault Detection (DFD) distinctly different and non standard and application of the centralized techniques to these problems is not straight forward.

However, some researchers [2,3] considered DFD problem with some kind of communication and made it possible to apply the algorithms developed for centralized situations The present authors, however, developed various approaches [4] to solve the DFD problem when there is no communication among the detectors. One of these techniques is the parity space approach [5].

The parity space approach uses analytic redundancy for FDI. One of the key aspects of analytic methods of FDI is the use of information concerning the dynamic relationships among the measured inputs and outputs. This approach considers residuals generated from a window of observations over some interval of time. Then the "observation space" has dimension proportional to the length of this interval, and at its simplest level, failure detection can be achieved by projecting the observations onto the orthogonal complement (parity space) of the subspace in which the observations should lie if ther is no failure.

However, this idea raises additional questions which are of importance in implementation Specifically, if we are given a linear, time-invariant system, how do we determine the minimum length of the window of observations that is required for detection and isolation? Can we generate all possible parity vectors from the given minimum order parity vectors? These questions are answered in a natural way considering the DFD problem formulation in the frequency domain in this paper. At the end we provide a numerical example to illustrate our methodology.

2. BACK-GROUND OF PARITY SPACE

Consider a linear, time-invariant, continuous time system described by

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad \ldots (1)$$

$$y(t) = Cx(t) \qquad \ldots (2)$$

where x,y,u are $n,m,\bar{q}$ dimensional vectors respectively. As a first step we will focus on sensor and plant failure detection and not actuator failures. Thus without loss of generality we will assume u=0. Therefore the system equations we are dealing with can be written as

$$\dot{x}(t) = Ax(t) \qquad \ldots (3)$$

$$y(t) = Cx(t) \qquad \ldots (4)$$

Define the extended observation vector of length p to be

$$\bar{y}_p(t) = [y^T(t), \dot{y}^T(t), \ldots, y^{(p-1)^T}(t)]^T \qquad \ldots (5)$$

obviously

$$\bar{y}_p(t) = \begin{bmatrix} C \\ CA \\ \cdot \\ \dot{C}A^{p-1} \end{bmatrix} x(t) \qquad \ldots (6)$$

Let q=mp, then $\bar{y}_p \in \mathbb{R}^q$. Define

$$\bar{C}_p = [C^T, (CA)^T, \ldots, (CA^{p-1})^T]^T \qquad \ldots (7)$$

and the observation space(of length p) $\mathbb{Z}_p$ to be

$$\mathbb{Z}_p = R(\bar{C}_p) \qquad \ldots (8)$$

where $R(\bar{C}_p)$ is the range of $\bar{C}_p$. Define the detection space (of length p) $\mathbb{G}_p$ to be the orthogonal compliment of $\mathbb{Z}_p$, i.e.

$$\mathbb{Z}_p \oplus \mathbb{G}_p = \mathbb{R}^q \qquad \ldots (9)$$

Then for any vector $g \in \mathbb{G}_p$, we have

$$\langle g, \bar{y}_p \rangle = 0 \qquad \ldots (10)$$

Equation (10) (or more precisely, the left hand side of this equation) is called a parity check. We will call g a detection vector (of length p).

3. FREQUENCY DOMAIN FORMULATION

In this section we establish the link between the time domain and frequency domain description of the parity checks. Then we show that the frequency domain description of all possible parity cheks is nothing more than the left nullspace of the polynomial matrix $[C^T,(sI-A)^T]^T$.

The observation space $\mathbb{Z}_p$ is a subspace spanned by the extended observation vector $\bar{y}_p(t)$ given in equn.(5). Assuming that there are no uncertainties, the detection space $\mathbb{G}_p$ is the orthogonal complement of $\mathbb{Z}_p$. We also define a detection vector a_p to be

$$a_p^T = [a_{p0}^T, a_{p1}^T, \dots, a_{pp-1}^T] \qquad \dots (11)$$

which satisfies

$$a_p^T \, \bar{y}_p(t) = 0 \qquad \dots (12)$$

for some $p = 0,1,\dots$. Recall that the relationship (12) is called a parity check. It is obvious that $a_p \in \mathbb{G}_p$ and finding all possible parity checks of length p is equivalent to finding a basis for detection space $\mathbb{G}_p$. Taking laplace transform of (12), we get

$$\sum_{i=0}^{p-1} a_{pi}^T \; s^i [y(s) - \sum_{j=0}^{i-1} y^{(j)}(0) s^{-(j+1)}]$$

$$= \sum_{i=0}^{p-1} s^i a_{pi}^T y(s) \; - \; \sum_{i=0}^{p-1} a_{pi}^T \; \sum_{j=0}^{i-1} y^{(j)}(0) s^{i-j-1}$$

$$= w^T(s) y(s) - \sum_{i=0}^{p-1} a_{pi}^T \sum_{j=0}^{i-1} y^{(j)}(0) s^{i-j-1} = 0 \qquad \dots (13)$$

But

$$y(s) = C(sI-A)^{-1} x_0 \qquad \dots (14)$$

$$y^{(p)}(0) = CA^p x_0 \qquad \dots (15)$$

Thus we have

$$w^T(s)C(sI-A)^{-1}x_0 = \sum_{i=0}^{p-1} a_{pi}^T \sum_{j=0}^{i-1} CA^j s^{i-j-1} x_o \qquad \ldots (16)$$

Since equn.12 is valid for all $x_0 \in \mathbb{R}^n$ we have

$$w^T(s)C(sI-A)^{-1} = \sum_{i=0}^{p-1} a_{pi}^T \sum_{j=0}^{i-1} CA^j s^{i-j-1} = r^T(s) \qquad \ldots (17)$$

Since equn.12 is nothing more than a parity check

$$w^T(s) = \sum_{i=0}^{p-1} a_{pi}^T s^i \qquad \ldots (18)$$

can be thought of as its frequency domain description, thus setting up a link between the frequency and time domains.

Now we will show that w(s) and r(s) span the left nullspace of the polynomial matrix $[C^T,(sI-A)^T]^T$. Here the left nullspace of some pxm rational matrix H(s) is defined as a subspace spanned by px1 rational vectors v(s)'s, which satisfy

$$v^T(s)H(s) = 0 \qquad \ldots (19)$$

Note that equn.17 can be written as

$$w^T(s)C = r^T(s)(sI-A)$$

or

$$[w^T(s), -r^T(s)]\begin{bmatrix} C \\ (sI-A) \end{bmatrix} = 0 \qquad \ldots (20)$$

Define $[w^T(s), -r^T(s)]^T$ as a detection vector and denote the left nullspace of $[C^T, (sI-A)^T]^T$ as N_{1c}. Then from equn.(20) we see that the polynomial parity vector$[w^T(s),-r^T(s)]^T \in N_{1c}$. Therefore it is enough to know one polynomial basis of N_{1c} in order to generate all parity checks of all possible finite lengths.

4. DECENTRALIZED FAULT DETECTION: PROBLEM FORMULATION

In this section, we formulate the decentralized fault detection problem. Consider the large interconnected system comprising of N subsystems described by

$$\dot{x}_i(t) = A_i x_i(t) + B_i u_i(t) + \sum_{\substack{j=1 \\ j\neq i}}^{N} A_{ij} x_j(t); \quad i=1,2,\ldots,N. \qquad \ldots (21)$$

$$y_i(t) = C_i x_i(t); \qquad i=1,2,\ldots,N. \qquad \ldots (22)$$

where $x_i(t)$ is an $n_i x1$ state vector, $u_i(t)$ is an $m_i x1$ input vector, $y_i(t)$ is a $p_i x1$ output vector of the i^{th} subsystem.

Here we are concerned with the problem of designing N - local detectors for fault detection. The local detectors donot communicate among themselves. We have, available at each local detector, the following:

1. measurements of the local input $u_i(t)$ and the local output $y_i(t)$,
2. the nominal values of the coefficient matrices C_i, A_i, B_i, A_{ij} and

In (21), if $A_{ij}=0$ or information about all x_i's is available at the i^{th} detector the problem is a classical one. Otherwise, each interaction term $A_{ij}x_j$ acts as an unknown input which makes the detection problem difficult to solve. By appropriately redefining various quantities, we can rewrite the term $\sum A_{ij}x_j$ as $E_i d_i(t)$ with E_i having full coulmn rank q_i in all the subsystems. Thus the i^{th}subsystem is

$$\dot{x}_i(t) = A_i x_i(t) + B_i u_i(t) + E_i d_i(t) \qquad \ldots (23)$$

$$y_i(t) = C_i x_i(t) \qquad \ldots (24)$$

We further assume that $p_i > q_i$, and $C_i E_i$ is also of rank q_i.

Now we formulate the frequency domain version of the DFD problem. Since d_i is unmeasurable and only measurements available are y_i and u_i the parity equation is given as

$$[\beta_i(s), -\alpha_i(s)] \begin{bmatrix} y_i(s) \\ B_i u_i(s) \end{bmatrix} = 0$$

or

$$[\beta_i(s), -\alpha_i(s)] \begin{bmatrix} C_i & 0 \\ (sI-A) & -E_i \end{bmatrix} = 0 \qquad \ldots (25)$$

with the condition $p_i > q_i$ we can always find $[\beta_i(s), -\alpha_i(s)]$, such that equation (25) is satisfied. Hence the detection vector in the frequency domain is $W^T(s) = (\bar{\beta}_i(s), -\bar{\alpha}_i(s))$, where '-' indicates the row vectors of β_i and α_i respectively. We rewrite the condition (25) to get two equivalent conditions as

$$\beta_i(s)C_i - \alpha_i(s)(sI-A) = 0 \qquad \ldots (26)$$

$$\alpha_i(s)E_i = 0 \qquad \ldots (27)$$

5. SOLUTION TO DECENTRALIZED DETECTION PROBLEM

We now consider the problem of obtaining the solution to (26) and (27). To this end, we go through the following sequence of steps to recast (26) and (27) in a matrix pencil form, from which the desired solutions can be easily obtained.

Let R_i denote the left annihilator of E_i, i.e.,

$$R_i E_i = 0 \qquad \ldots (28)$$

where R_i has maximal row-rank. From (27), we have then

$$\alpha_i(s) = \tilde{\alpha}_i(s) R_i \qquad \ldots (29)$$

for some $\tilde{\alpha}_i(s)$. Then solving $\tilde{\alpha}_i(s)$ is equivalent to finding α_i, since rank α_i = rank $\tilde{\alpha}_i$ (this follows from applying Sylvester's inequality to (29) and using the fact that R_i has full row rank). Substituting (29) in (26), we get

$$\tilde{\alpha}_i(s) R_i (sI-A) = \beta_i(s) C_i \qquad \ldots (30)$$

Let Q_i = Ker C_i, i.e. $C_iQ_i = 0$, where Q_i has maximal column rank and let C_{iR} be the right inverse of C_i, i.e. $C_iC_{iR} = I$. Post multiplying (30) by Q_i and C_{iR} respectively, we get

$$\tilde{\alpha}_i(s)(sR_iQ_i - R_iA_iQ_i) = 0 \qquad \ldots (31)$$

$$\tilde{\alpha}_iR_i(sI-A_i)C_{iR} = \beta_i(s) \qquad \ldots (32)$$

So, once $\tilde{\alpha}_i$ is obtained from (31), β_i can be directly computed using (32). Since $(sR_iQ_i - R_iA_iQ_i)$ is a singular pencil, the problem thus, reduces to obtaining the left null-space of the matrix pencil. A simple way of obtaining the left null-space $(sR_iQ_i - R_iA_iQ_i)$ is by transforming it to the Kronecker canonical form and use the well established results on this form for finding the left null-space. From Kronecker's theory, there exists non-singular square matrices M and N such that the Kronecker canonical form can be obtained as

$$M_i(sR_iQ_i - R_iA_iQ_i)N_i = \text{Block diag. } (O_{gxh}, L_e, L_r^T, D_a, (sI-J)) \quad \ldots (33)$$

where L_e, L_r^T, D_a and $(sI-J)$ have diagonal forms. Detailed discussion on this canonical form may be found in Gantmacher [6].

Now we show how in a simple way vectors of the left null-space of a matrix pencil can be obtained when it is in the Kronecker canonical form. Consider the right hand side of (33). We note that in writing the above canonical form, we assumed that the pencil has minimal column indices $\{e_1, e_2, \ldots, e_\eta\}$ and minimal row indices $\{r_1, r_2, \ldots, r_t\}$. The D_a and J blocks correspond to finite and infinite elementary divisors respectively.

It is easy to see that the left null - space of an operator with a block diagonal matrix representation is the direct sum of the left null - spaces of the operators corresponding to each of the diagonal blocks.

Let U = block diag. $(\Delta, \bar{P}, P, \Theta, Z)$ ... (34)

where Δ, $\bar{P}$, P, Θ and Z are the collection of vectors of the null-spaces corresponding to O_{gxh}, L_e, L_r^T, D_a and $(sI-J)$ blocks respectively. Note further, that in view of the structure of (33), $\bar{P}$, P, D and Z if they exist, are themselves block diagonal. Now we consider the coponents of U. From the dimensions of O_{gxh} block, it is evident that

$$\Delta = I_g \qquad \text{... (35)}$$

Each of the components L_{e_i} of L_e has full row rank for all values of s and hence L_e has full row rank for all s. Consequently, $\bar{P}$ has the trivial solution

$$\bar{P} = 0 \qquad \text{... (36)}$$

We have P = Block diag. $(P_{g+1},\dots,P_t)$, where each P_{g+i} corresponds to the null-space of $L^T_{r_{g+i}}$. The equation

$$\bar{p}_i(s)L^T_{r_{g+i}} = 0 \qquad \text{... (37)}$$

has a polynomial solution of the form

$$\bar{p}_i(s) = (1, s, \dots, s^{r_{g+i}}) \qquad \text{... (38)}$$

As we are interested only in this solution, we donot consider other solutions of P. Next, since D_{a_i} (and hence D_a) has full row rank for all s, we have

$$\Theta = 0 \qquad \text{... (39)}$$

Finally we are interested in the solution of

$$z_{ij}\,(sI-J_{ij}) = 0 \qquad \text{... (40)}$$

Observe that $s \neq \lambda_i$, where λ_i's are the eigenvalues associated with J_{ij} and $(sI-J_{ij})$ is nonsingular and hence (40) has only the trivial solution $z_{ij} = 0$. Thus, finally we have U as

$$U = \text{block diag.}\ (I_g,\ O,\ P,\ O,\ O) \qquad \ldots (41)$$

Observe that the solution U obtained above is precisely the solution that we want.

Defining

$$U_i = \tilde{\alpha}_i M_i^{-1} \qquad \ldots (42)$$

(31) can be put into the form (33) as

$$U_i(sM_i \cdot {}_iQ_iN_i - M_iR_iA_iQ_iN_i) = 0 \qquad \ldots (43)$$

If U_i is determined from (43), α_i can be computed from (29). Thus the original problem has reduced to one of finding U_i satisfying (43) and this U_i is nothing more than that given by (41).

6. NUMERICAL EXAMPLE

To illustrate the technique described in the previous section here we consider a simple numerical example. Consider the observable system with A_i, B_i, C_i and E_j given by

$$A_i = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & -1 \end{bmatrix};\ B_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix};\ C_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix};\ E_i = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

Observe that $p_i = 2$, $q_i = 1$, i.e. $p_i > q_i$. By direct calculation it is true that $Q_i = \text{Ker}\ C_i = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$; $R_i = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $R_iQ_i = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$; $R_iA_iQ_i = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ $(sR_iQ_i - R_iA_iQ_i) = \begin{bmatrix} s \\ 1 \end{bmatrix}$. Thus, with $M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, the Kronecker form for the pencil is given by $\begin{bmatrix} s \\ -1 \end{bmatrix}$. Thus the pencil has only one block of the type L_r^T with $r = 1$. We note that

$$\bar{P}_1(s)L^T_{r_i}(s) = [1,\ s]\begin{bmatrix} s \\ -1 \end{bmatrix} = 0$$

Thus $P_i = [1,\ s]$, from which we get

$\alpha_i(s) = [0,\ 1,\ -s]$ and $\beta_i(s) = [1,\ -s-s^2]$

So the parity equation in frequency domain is

$$[1,\ -s-s^2,\ 0,\ -1,\ s]\begin{bmatrix} y_i(s) \\ B_i u_i(s) \end{bmatrix} = 0$$

7. CONCLUSIONS

In this paper we have given a frequency domain formulation for the DFD problem and established how naturally the questions on minimal order parity vector and minimum length of the window of observations can be answered.

8. REFERENCES

1. Willsky, A.S., Automatica, Vol.12, pp.601-611, 1976.
2. Singh, M.G. et al, Porc. IEE, Vol.130, pp. 243-249, 1983.
3. Hodzic, M. et al, Large Scale Systems, Vol.9, pp.83-100, 1985.
4. Gopala Rao, K.A., Ph.D. Thesis, I.I.Sc., Bangalore, 1986.
5. Chow, E.Y. et al, Proc. IEEE conf. on Decision and Control, pp.1006 -1012, Dec. 1980.
6. Gantmacher, F.R., The Theory of Matrices, Vol.I & II, 1959.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 167-176

SENSOR FAULT DETECTION WITH REDUCED ORDER MODELS

K.A. GOPALA RAO
DEPT. OF ELECTRICAL ENGINEERING
ANDHRA UNIVERSITY
VISAKHAPATNAM, INDIA.

N. VISVANADHAM
DEPT. OF COMPUTER SCIENCE & AUTOMATION
INDIAN INSTITUTE OF SCIENCE
BANGALORE, INDIA.

ABSTRACT

In this paper, we consider fault detection and diagnosis in sensor of a singularly perturbed system using single observer scheme of Clark. We design an observer based on reduced order model of the system obtained by neglecting the fast dynamics. Our main concern here is to study the effects of neglected fast dynamics on fault detection and diagnosis We show here that the modlling errors due to neglected fast dynamics generate non-zero residuals even under the conditions of normal operation and from the expressions for the bounds on these terms we show that these terms are in general small in magnitude. Thus the standard practice of control engineering of using reduced models is shown to apply equally well to fault detection and diagnosis problems. However if in a particular problem one feels that even this contribution in the normal operation may lead to false alarms in fault detection and diagnosis, then one can design time-varying alarm thresholds based on the bounds given for these terms.

INTRODUCTION

Systems with small parameters are common in control problems. These small parameters, with values proportional to a small positive number, represent weak connections are parasitics. Since they appear as coefficients in the dynamical equations representing the system, these parasitics increase the order of the system and also make the differential or difference equations stiff. The curse of the dimensionality coupled with stiffness poses formidable computational difficulties for the analysis and control of such systems.

Use of lower order models to reduce computational and implementation complexities with small performance degradation is a standard strategy in control theory and practice. Orde reduction based on singular perturbation method is a well developed methodology [1,2]. However, no results have been reported so far in the field of Fault Detection and Dignosis (FDD) using this approach.

In this paper, we study the effects of order reduction on FDD based on observers. The paper is organized as follows. We first present some back ground material on singularly perturbed systems. We then consider the fault detection problem in singularly perturbed systems based on the single observer scheme of Clark [3]. Here, we explicitly find the errors in estimation due to the use of reduced order model under normal conditions and derive expressions for bounds needed in selecting thresholds for FDD of sensor faults.

PRELIMINARIES

Consider a singularly perturbed linear time-invariant system described by

$$\dot{x}_1(t) = A_{11}x_1(t) + A_{12}x_2(t) + B_1u(t);\ x_1(0) = x_{10} \qquad \ldots (1)$$

$$\mu\dot{x}_2(t) = A_{21}x_1(t) + A_{22}x_2(t) + B_2u(t);\ x_2(0) = x_{20} \qquad \ldots (2)$$

$$y(t) = C_1x_1(t) + C_2x_2(t) \qquad \ldots (3)$$

where μ is a small positive scalar, the state vectors x_1 and x_2 are respectively of dimension n_1 and n_2, the input u is of dimension m and the output y is of dimension p. Equation (1) is generally known as slow system dynamics. In general if we are given the dynamics of the plant to be controlled and the dynamics of the fast actuators and sensors, then in (1) - (3) x_1 represents the state of the plant and x_2 represents the state of the fast actuator and sensor dynamics. Now neglecting parasitics is equivalent to setting $\mu = 0$ in system (1) - (3). The resulting system is known as the reduced system and is given as (assuming A_{22} is nonsingular)

$$\dot{\bar{x}}_1(t) = A_0\bar{x}_1(t) + B_0u(t); \; \bar{x}_1(0) = x_{10} \qquad \ldots (4)$$

$$\bar{x}_2(t) = -A_{22}^{-1}[A_{21}\bar{x}_1(t) + B_2u(t)] \qquad \ldots (5)$$

$$\bar{y}(t) = C_0\bar{x}_1(t) + D_0u(t) \qquad \ldots (6)$$

Equation (4) is known as the slow reduced system and

$$A_0 = A_{11} - A_{12}A_{22}^{-1}A_{21} \qquad \ldots (7)$$

$$B_0 = B_1 - A_{12}A_{22}^{-1}B_2 \qquad \ldots (8)$$

$$C_0 = C_1 - C_2A_{22}^{-1}A_{21} \qquad \ldots (9)$$

$$D_0 = -C_2A_{22}^{-1}B_2 \qquad \ldots (10)$$

It is well known [4,5] that if the pair (C_0, A_0) is observable then it is possible to construct a full-order observer of the form

$$\dot{\hat{\bar{x}}}_1(t) = (A_0 - GC_0)\hat{\bar{x}}_1(t) + (B_0 - GD_0)u(t) + G\bar{y}(t); \; \hat{\bar{x}}_1(0)=\bar{x}_1(0) \qquad \ldots (11)$$

where the matrix G is chosen to place the poles of $(A_0 - GC_0)$ so that (11) is a stable observer of the slow reduced system, i. e. the error

$$\bar{\varepsilon}(t) = \bar{x}_1(t) - \hat{\bar{x}}_1(t) \qquad \ldots (12)$$

which satisfies

$$\dot{\bar{\varepsilon}}_1(t) = (A_0 - GC_0)\bar{\varepsilon}_1(t) \qquad \ldots (13)$$

will decay exponentially from any initial error $\bar{\varepsilon}_1(0) = \bar{x}_1(0) - \hat{\bar{x}}_1(0)$. However if the observer (11) is used to observe the slow state of the system (1) it will have exponential decay term due to (13) and also other error terms caused by neglecting fast dynamics (2). Javid [6] has derived the form of this error term and the time-varing upper bound for this error term. Since we require this in the further development we reproduce the same here.

Let u(t) be continuous and bounded together with its first derivative for $t \geq 0$.

Now define

$$\varepsilon_1(t) = x_1(t) - \hat{\bar{x}}_1(t) \qquad \ldots (14)$$

then

$$\varepsilon_1(t) = e^{(A_0-GC_0)t}\,\varepsilon_1(0) + U(t,\mu) \qquad \ldots (15)$$

where $U(t,\mu)$ represents the error due to the neglected fast dynamics and the bound on this is defined as

$$\|U(t,\mu)\| \leq \mu\,\bar{k}(0)[e^{-\alpha_0 t} - e^{-\alpha t/\mu}] \qquad \ldots (16)$$

with

$$\bar{k}(0) = \frac{k_0 k}{\alpha - \mu\alpha_0}\,\|A_{12} - GC_2\|\,\|x_2(0) + A_{22}^{-1}A_{21}x_1(0) + A_{22}^{-1}B_2u(0)\| \qquad \ldots (17)$$

k_0, α_0, k and α are chosen as discussed in [6]. In the above equations $\|\cdot\|$ indicates the matrix norm which is defined as sum of the absolute values of the elements of the matrix refered in the norm. It can be seen from (16) that this error is fortunately of order μ.

With this background now we discuss in the next section sensor fault detection and diagnosis using the observer (11).

OBSERVER BASED FAULT DETECTION AND DIAGNOSIS

In this section we consider FDD in sensors of a singularly perturbed system using single observer Scheme of Clark [3]. We design an observer based on the reduced order model of the system obtained by neglecting the fast dynamics [6]. Our main concern here is to study the effects of neglected fast dynamics on FDD. We show here that the modelling errors due to neglected fast dynamics generate nonzero residuals even under the conditions of normal operation and from the expressions for the bounds on these terms we show that these terms are in general small in magnitude. Thus the standard practice of control engineering of using reduced models is shown to apply equally well to FDD problems. However if in a paticular problem one feels that even this contribution in the normal operation may lead to false alarms in FDD, then one can design time-varying thresholds based on the bounds given for these terms.

Clark and his coworkers [3,7,8] developed two observer schemes for sensor fault detection: (i) dedicated observer sheme, (ii) single observer scheme. In the dedicated observer scheme a bank of observers is used for FDD and each sensor will drive a single observer. In the single observer scheme FDD is done with the estimates obtained from a single observer.

Here we cosider the single observer scheme of Clark [3]. In this method the state estimates are obtained using a single observer driven by the output of a single sensor and the inputs to the original system. We make the following assumptions:

1. c_{1i} and c_{2i} represent the i^{th} rows of C_1 and C_2 respectively.

2. The pair (A_{22}, c_{2i}) and the pair (A_0, C_0^i), where C_0^i is C_0 obtained

with c_{1i} and c_{2i}, are observable.

3. A single sensor is faulty at a time.

The method of sensor fault detection is as follows. From the state estimates, we estimate the output of each sensor. Then we compare these output estimates with the actual outputs fromthe sensors as follows:

$$v_j = y_j - \hat{y}_j \qquad \ldots (18)$$

where the subscript j indicates the j^{th} sensor.

Under normal operating mode without any modelling errors, $v_j=0$ and if there is any fault v_j will have a large non-zero value. But to avoid the effects of noise, modelling error etc., a threshold is used for FDD on v_j. Now in the present problem we have a singularly perturbed system and we are estimating the states of the original system using an observer of the type (11). We have seen that in this case the estimation error $\varepsilon_1(t)$ is non-zero even under normal conditions due to $U(t,\mu)$ and hence v_j is non-zero even under normal conditions. If these errors have substantial contribution then they may lead to false alarms if the thresholds are fixed without taking into consideration these error terms. Since we know from [6] that we can explicitly find the bounds on these terms, these bounds will be useful in fixing the thresholds for fault detection so as to minimise the false alarms due to neglected fast dynamics and ensure reliable alarming when sensor faults occur. We derive the bounds on the errors in v_j due to neglected fast dynamics in this section.

The observer driven by the i^{th} sensor is

$$\dot{\hat{\bar{x}}}_1(t) = (A_0 - GC_0^i)\hat{\bar{x}}_1(t) + (B_0 - GD_0^i)u(t) + Gy_i(t) \qquad \ldots (19)$$

and the estimate of x_2 is

$$\hat{x}_2(t) = -A_{22}^{-1}\,[A_{21}\hat{x}_1 + B_2u(t)] \qquad \ldots (20)$$

Let $\varepsilon_1(t)$ and $\varepsilon_2(t)$ be the estimation errors corresponding to x_1 and x_2 then they are given as

$$\varepsilon_1(t) = e^{(A_0-GC_0^1)t}\;\varepsilon_1(0) + U_1(t,\mu) \qquad \ldots (21)$$

$$\varepsilon_2(t) = -A_{22}^{-1}\,A_{21}e^{(A_0-GC_0^1)t}\;\varepsilon_1(0) + U_2(t,\mu) \qquad \ldots (22)$$

where U_1 and U_2 are the error terms due to the neglected fast dynamics and the bounds

$$\|U_1(t,\mu)\| \leq \mu\;\bar{k}_1(0)[e^{-\alpha_0 t} - e^{-\alpha t/\mu}] \qquad \ldots (23)$$

$$\|U_2(t,\mu)\| \leq \mu\;\bar{k}_2(0)[e^{-\alpha_0 t} - e^{-\alpha t/\mu}] \qquad \ldots (24)$$

where $\bar{k}_1(0)$ is given by (17) and

$$\bar{k}_2(0) = \frac{k_0k}{\alpha-\mu\alpha_0}\;\|A_{22}^{-1}\|\;\|A_{21}\|\;\|A_{12}-GC_2\|\;\|x_2(0)+A_{22}^{-1}[A_{21}x_1(0)+B_2u(0)]\| \qquad (25)$$

similarly we can find the bounds on the errors in v_i.

Let $\bar{U}^i(t,\mu)$ error term corresponding to v_i then the bound on $\bar{U}^i$ is

$$\|\bar{U}^i(t,\mu)\| \leq \|c_{1i}\|\;\|U_1(t,\mu)\| + \|c_{2i}\|\;\|U_2(t,\mu)\| \qquad \ldots (26)$$

The significance of these bounds is that, even under conditions of normal operation, the estimation errors and hence v_i will not be zero but have some finite values with upper bounds given above. This happens because of the use of reduced system observer with the fast dynamics neglected. However these errors are of order μ, hence the performance of FDD using this observer will not be degraded too much if one uses the observer (11) for FDD. However if one feels that the contribution due to these error terms is substantial then if the alarm thresholds are designed based on these upper bounds of the error terms, false alarms due to the neglected fast dynamics can be minimized and ensure reliable alarming when sensor faults occur.

EXAMPLE

Consider the following singularly ~~perturbed~~ system:

$$\dot{x}_1(t) = -x_1(t)$$

$$\mu\dot{x}_2(t) = x_1(t) - x_2(t)$$

$$y_1(t) = x_1(t) + x_2(t)$$

$$y_2(t) = 2x_1(t) + x_2(t)$$

With $\mu = 0$, we have

$$x_1(t) = e^{-t}\, x_1(0)$$

$$x_2(t) = e^{-t/\mu}\, x_2(0) + \frac{[e^{-t} - e^{-t/\mu}]}{(1 - \mu)}\, x_1(0)$$

We wish to construct an observer using only y_1 and hence the reduced system in which we are interested is

$$\dot{\bar{x}}_1(t) = -\bar{x}_1(t)$$

$$\bar{x}_2(t) = \bar{x}_1(t)$$

$$\bar{y}_1(t) = 2\bar{x}_1(t)$$

The observer for the reduced system is

$$\dot{\hat{\bar{x}}}_1(t) = -(1 + 2g)\,\hat{\bar{x}}_1(t) + g\bar{y}_1(t)$$

Then with $g=2$ we have

$$\hat{\bar{x}}_1(t)=e^{-5t}\hat{\bar{x}}_1(0)+[\frac{19}{9}\, e^{-t}- \frac{3}{2}\, e^{-5t}+ \frac{4}{9}\, e^{-10t}]\bar{x}_1(0)+ \frac{2}{5}[e^{-5t}-e^{-10t}]\bar{x}_2(0)$$

We have chosen the initial conditions $x_1(0)=x_2(0)=\bar{x}_1(0)=1$, $k=1$, $\alpha=1=\alpha_0$, $k_0=0.0625$ and found the bounds on output error $\Omega^{\cdot}$ and Ω^2, and compared the same with actual Ω^1 and Ω^2 as shown in the following table.

time sec	error in y_1-y_1	bound on this error (Ω^1)	error in y_2-y_2	bound on this error (Ω^2)
0	0.333	0	0.666	0
0.1	0.34935	0.298311	0.645	0.4475
0.2	0.31942	0.3797	0.5776	0.5695
0.3	0.2816	0.3839	0.5075	0.57586
0.4	0.2468	0.36222	0.4458	0.54334
0.5	0.21725	0.33322	0.3936	0.499827
1	0.12396	0.2043776	0.2268	0.30657
5	2.24598×10^{-3}	3.7433×10^{-3}	4.1176×10^{-3}	5.61495×10^{-3}
10	1.51332×10^{-5}	3.22843×10^{-5}	2.7744×10^{-5}	4.8426399×10^{-5}

In the observer based techniques the FDD mechanism is initiated in the steady state region of the observer. The error bounds (17) and (25) indicate that as $t \to \infty$ they go to zero. Considering these two facts one can design the alarm thresholds as follows: If the decay of the error bounds is not fast compared to the observer then the error terms have substantial magnitude and thus to avoid false alarms one has to fix a threshold which is time-varying, based on the error bounds (17) and (25). However if the error terms decay fast then the usual method of constant threshold can be adopted. The above example illustrates that the effects of neglected fast dynamics is negligible in the steady state region of the observer (settling time for observer is 5 sec) and hence a constant alarm threshold can be designed.

CONCLUSIONS

In this paper, we have shown that the standard practice in control engineering of using lower order models to reduce the computational burden with less performance degradation applies equally well to the problem of fault detection and diagnosis. In particular we have studied the following aspects. Using the observed based technique we have shown that when an observer of the reduced system is used for FDD, the contribution of the neglected fast dynamics to the estimation error is a term of order μ and hence there will not be much performance degradation. However if in a particular problem this contribution is considered to be substantial and causes false alarms, then one can design the time-varying alarm thresholds based on the expressions given for the bounds on these error terms.

REFERENCES

1. P.V. Kokotovic, R.E. O'Malley and P. Sannuti, "Singular perturbation and order reduction in contol theory: an overview", Automatica, Vol. 12, pp. 123-132, 1976.
2. V.R. Saksena, J. O,Reilly and P.V. Kokotovic, "Singular perturbation and time-scale methods in control theory: Survey 1976 - 1983", Automatica, Vol. 20, pp. 273-293, 1984.
3. R.N. Clark, "A simplified instrument failure detection scheme", IEEE Trans. Aero. and Electronic systems, Vol. 14, pp. 558-563, 1978.
4. D.G. Luenberger, "Observing the state of a linear system", IEEE Tran Military Eletronics, Vol. 8, pp. 74-80, 1964.
5. D.G. Luenberger, "An introduction to observers", IEEE Trans. Automatic Control, Vol. 16, pp. 596-602, 1971.
6. S. H. Javid, " Observing the slow states of a singularly perturbed systems", IEEE Trans. Automatic Control, Vol. 25, pp. 277-280, 1980.
7. R.N. Clark, D.C. Fosth and V.M. Walton, "Detecting instrument malfunctions in control systems", IEEE Trans. Aero. and Electronic systems", Vol. 11, pp. 465-473, 1975.
8. R. N. Clark, " Instrument fault detection ", IEEE Trans. Aero. and Electronic systems, Vol. 14, pp. 456-465, 1978.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 177-185

CONTROL OF A SYNCHRONOUS GENERATOR OVER WIDE RANGE OF OPERATING CONDITIONS BY EIGENVALUE PLACEMENT

K.A. GOPALA RAO

DEPT. OF ELECTRICAL ENGINEERING

ANDHRA UNIVERSITY, VISAKHAPATNAM

INDIA

ABSTRACT

In this paper it is shown that the power system stabilizers based on the state feedback technique can be designed using standard forms instead of a polynomial corresponding to an arbitrary dominant second order cofiguration [1]. It is also shown that substantial performance insensitivity to operating point can be achieved, i.e. state feedback controllers designed by the eigenvalue placement technique are suitable for controlling, over wide range of operating conditions, a synchronous generator connected to a large system. A simplified version of the phase variable feedback controller is also investigated.

INTRODUCTION

The importance of excitation control [2] in improving the dynamic stability of synchronous generators is widely recognized. Much effort has been devoted to apply optimal control theory to systematic design of exciter and/or prime-mover control stratagies, for example [3 - 5]. Later, the technique of eigenvalue placement (or model control) has been applied [1, 6, 7] to this problem. The eigenvalue placement technique rests on the important result of linear control theory, namely, that the eigenvalues of closed - loop system can be arbitrarily located by linear state feedback, provided the given system is controllable[8]. The computation of the required feedback coefficients is systematic and at the same time non-iterative [1], compared to optimal control design. Thus, an excitation controller can be designed [1] wich gives arbitrary desired system eigenvalues when generator is operating at a nominal operating point.

However, in most power system controller design the parameters of the controller are chosen in such a way that the resulting closed-loop system has dominant second order configuration. Many times this requires a high control effort which is unacceptable for practical implementation. However, in control theory standard forms are developed which give desired eigen patterns and at the same time require less control effort.

The present paper deals with the suitability of these models in the design of controllers for wide range of operating points in the case of a generator connected to a large system. It is also shown that substantial performance insensitivity to operating point can be achieved with these controllers. A simplified version of the phase variable feedback controller is also studied. It is shown that the performance of the controller is satisfactory over a wide range of operating conditions.

PROPOSED METHOD

The small perturbation block diagram model of a synchronous machine infinite bus system of [2] is considered for the analysis and design of controllers. To apply the eigenvalue placement technique the system equations are written in the phase variable form as:

$$\dot{x}(t) = A\ x(t) + b\ u(t) \qquad \ldots (1)$$

where the state vector x(t) is given by

$$x^T(t) = [x_1, x_2, x_3, x_4]^T$$

$$= [\Delta\delta, \Delta\dot{\delta}, \Delta\ddot{\delta}, \Delta\dddot{\delta}]^T \qquad \ldots (2)$$

u(t) is the supplementary control input to be designed using state feedback. The A and b constant matrices for the system are given in appendix. In the eigenvalue placement technique, a linear state variable feedback control law

$$u(t) = -F\,x(t) \qquad \ldots (3)$$

is sought, which locates the eigenvalues of the closed-loop system:

$$\dot{x}(t) = (A - bF)\,x(t) \qquad \ldots (4)$$

at desired locations.

The special properties of linear phase variable feedback can be appreciated by considering the block diagram (Fig.1) of the system with phase variable feedback controller.

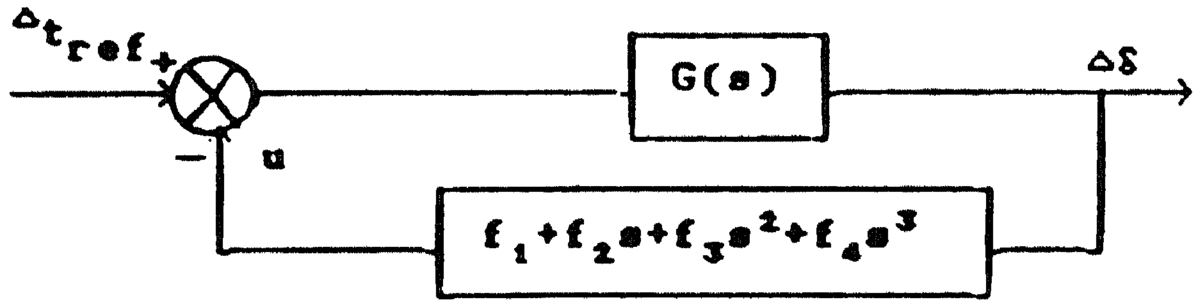


Fig. 1 Block Diagram Of The System With Phase Variable Feedback Controller

The transfer function G(s) is given in the appendix. The effect of-phase variable feedback is seen to be equivalent to introduction of a feedback compensator with zeros given by

$$f_1 + f_2s + f_3s^2 + f_4s^3 = 0 \qquad \ldots (5)$$

It may be noted from (5) that if the zeros are given, the relative values of the coefficients f_1 to f_4 are fixed, i.e. f_1 to f_4 are fixed to with in a constant multiplier. From root locus [8], it can be inferred that for large values of gain in the loop, three of the four eigenvalues of the closed-loop system will be close to the zeros of (5) and the fourth eigenvalue will be far away on the negative real axis of the s-plane. This will be true irrespective of the locations of the open-loop eigenvalues or poles of G(s) which change with the operating point Coupled with the observation that there is already a fairly large gain K_e in G(s), this suggests that the eigenvalues of the closedloop system of Fig. 1 can be made insensitive to the operating point with moderate feedback path gains f_1 to f_4.

In [1], a dominant pair eigenvalue configuration was suggested for the desired eigenvalues so that the response characteristics such as degree of stability (indicated by damping ratio ξ of the dominant pair) and speed of response (indicated by natural frequency w_n of the dominant pair) can be readily assessed. However, many times this requires a high control effort which is unacceptable for practical implementation. In control theory standard forms are developed which give desired eigen patterns and at the same time require less control effort.

To evaluate the suitability of eigenvalue placement controller designed based on standard forms for wide range of operating points, we consider a typical example [2] of a generator infinite bus system with the data:

$x_d = 1.6$, $x_d' = 0.32$, $x_q = 1.55$, $T_{d0}' = 6$

$H = 5$, $x_e = 0.4$, $T_e = 0.05$, $K_e = 200$

The range [2] of machine loadings considered is:

1 + j0, 1 + j0.5, 1 - j0.5, 0.5 + j0, 0.1 + j0.

The three zeros of (5) are chosen using different standard forms and the resulting coefficients are shown in Table - I. The eigenvalues for various operating conditions with and without these controllers are given in Tables II, III, and IV. With feedback coefficients increased by a factor of four, the closed - loop system eigenvalues for different operating conditions are in Tables V and VI.

TABLE I: FEEDBACK COEFFICIENTS

Polinomial	f_1	f_2	f_3	f_4
Butterworth	-2.6	-0.36	-0.03	-0.00125
ITAE	-2.6	-0.39	-0.0263	-0.00125

TABLE II EIGENVALUES WITHOUT CONTROLLER

Operating point	eigenvalues
(1) 1 + j0	0.78 ± j07.93 -11.01 ± j10.45
(2) 1 + j0.5	0.35 ± j6.78 -10.58 ± j13.52
(3) 1 - j0.5	2.84 ± j8.10 -13.07 ± j3.36
(4)0.5 + j0	-0.19 ± j5.95 -10.04 ± j14.00
(5)0.1 + j0	-0.06 ± j4.30 -10.17 ± j16.35

TABLE III EIGENVALUES WITH CONTROLLER (BUTTERWORTH DESIGN)

Operating point	Eigenvalues	ξ	ω_n
(1) 1 + j0	-4.68 ± j10.61 -17.30 , -39.91	0.40	11.6
(2) 1 + j0.5	-4.28 ± j09.78 -24.00 ± j09.42	0.40	10.7
(3) 1 - j0.5	-4.43 ± j11.10 -12.43 , -53.50	0.35	12.7
(4)0.5 + j0	-4.41 ± j09.25 -22.63 ± j11.32	0.43	10.3
(5)0.1 + j0	-1.77 ± j06.68 -12.67 ± j17.31	0.26	6.9

TABLE IV EIGENVALUES WITH CONTROLLER (ITAE DESIGN)

Operatin point	Eigenvalues	ζ	ω_n
(1) 1 + j0	-4.63 ± j12.68 -11.02 , -46.31	0.34	13.5
(2) 1 + j0.5	-4.69 ± j11.64 -15.00 , -32.30	0.37	12.6
(3) 1 - j0.5	-3.85 ± j13.73 -9.10 , -58.00	0.27	14.3
(4)0.5 + j0	-5.10 ± j11.15 -16.15 , -27.80	0.41	12.2
(5)0.1 + j0	-2.10 ± j06.95 -12.38 ± j16.30	0.28	7.3

TABLE V EIGENVALUES WITH FEEDBACK COEFFICIENTS MULTIPLIED BY FOUR (BUTTERWORTH DESIGN)

Operating point	Eigenvalues	ζ	ω_n
(1) 1 + j0	-5.67 ± j10.50 -13.00 ,-184.20	0.48	11.9
(2) 1 + j0.5	-5.54 ± j10.21 -13.94 ,-139.83	0.48	11.6
(3) 1 - j0.5	-5.69 ± j10.84 -12.11 ,-214.40	0.47	12.2
(4)0.5 + j0	-5.60 ± j10.10 -14.10 ,-129.70	0.48	11.5
(5)0.1 + j0	-4.10 ± j08.50 -23.00 ± j14.54	0.43	9.4

TABLE VI EIGENVALUES WITH FEEDBACK COEFFICIENTS MULTIPLIED BY FOUR (ITAE DESIGN)

Operating point	Eigenvalues	ξ	ω_n
(1) 1 + j0	-5.85 ± j12.84 -9.10 ,-184.20	0.42	14.1
(2) 1 + j0.5	-5.86 ± j12.60 -9.52 ,-143.60	0.42	13.9
(3) 1 - j0.5	-5.70 ± j13.20 -8.66 ,-217.80	0.40	14.4
(4)0.5 + j0	-5.96 ± j12.45 -9.50 ,-133.50	0.43	13.8
(5)0.1 + j0	-5.10 ± j09.87 -21.97 ± j06.97	0.46	11.1

It is noted from the tables V and VI that substantial insensitivity to operating point has been achieved. More insensitivity can be secured by further increasing f_1 to f_4 uniformly. This provides a quantitative approach to the problem of performance sensitivity to operating point.

SIMPLIFIED VERSION OF THE FEEDBACK CONTROLLER

From a practical view point, it is of interest to consider whether the $\dddot{\delta}$ term in the phase variable feedback controller can be dispensed with. By neglecting the time constant for the thyristor exciter in the system, a model with only three phase variables $\Delta\delta$, $\Delta\dot{\delta}$, $\Delta\ddot{\delta}$ can be obtained. By choosing the zeros of

$$f_1 + f_2 s + f_3 s^2 = 0 \qquad \ldots (7)$$

based on the second order polynomial of standard forms the feedback coefficients will be:

$f_1=-3.6$, $f_2=-0.84$, $f_3=-0.1$

for both the standard forms (since the second order polynomial is same for both the standard forms). The eigenvalues for different operating points with this controller are given in Table VII.

TABLE VII EIGENVALUES WITH SIMPLIFIED CONTROLLER

Operating point	Eigenvalues	ζ	ω_n
(1) 1 + j0	-4.13 ± j04.62 -6.10 ± j61.40	0.67	6.2
(2) 1 + j0.5	-3.99 ± j04.72 -6.24 ± j55.10	0.64	6.3
(3) 1 - j0.5	-4.28 ± j04.59 -5.95 ± j65.41	0.68	6.3
(4)0.5 + j0	-3.95 ± j04.59 -6.28 ± j53.40	0.65	6.1
(5)0.1 + j0	-2.94 ± j04.97 -7.29 ± j29.95	0.51	5.8

It is noted from table VII that the insensitivity property is largely retained, though the dominance of the primary pair of eigenvalues over the others is reduced as compared to the complete phase variable feedback controller. This simplified controller can still be considered as satisfactory over wide range of operating points.

CONCLUSIONS

It has been shown that state variable feedback excitation controllers designed by the eigenvalue placement technique are suitable for controlling over wide range of operating points a synchronous generator connected to a large system. Standard forms developed in control theory are used to choose feedback coefficients instead of the dominant pair eigenvalue concept. Substantial performance insensitivity to operating point has been achieved with these controllers. A simplified version of the feedback controller has also been investigated. Although the eigenvalue placement using state feedback approach to controller design for synchronous machines is not new, the method of selection of the feedback coefficients is believed to be new and yields a quantitative approach for securing performance insensitivity to operating point.

APPENDIX

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dfrac{-377K_eK_2}{MT'_{d0}T_e} \end{bmatrix}$$

$$G(s) = \frac{-377\ K_eK_2/MT'_{d0}T_e}{s^4 + a_3s^3 + a_2s^2 + a_1s + a_0}$$

where all coefficients are taken as defined in [1], and the constants are taken as in [2].

REFERENCES

1. M.A. Mohan et al, Paper no. A 78 249-5, IEEE PES winter meeting,1978
2. F.P. Demello and C. Concordia, IEEE Trans. PAS-88, 316-329, 1969.
3. Y.N. Yu et al, IEEE Trans. PAS-89, 55-62, 1970.
4. J.H. Anderson, Proc. IEEE, Vol. 59, 25-35, 1971.
5. H.A.M. Moussa and Y.N. Yu, IEEE Trans. PAS-91, 1166-1174, 1972.
6. H. Seraji, Proc. IEE, Vol.121, 529-532, 1974.
7. O.P. Malik and U.C. Chen, paper no. A 78 254-5, IEEE PES winter meeting, Jan. 1978.
8. J.J. D'azzo and C.H. Houpis, "Linear Control System Analysis and Design", McGraw-Hill Book Co., New York, 1975.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 187-192

DESIGN OF OPTIMAL DISCRETE CONTROL SYSTEMS WITH PRESCRIBED DEGREE OF STABILITY

K.A.GOPALA RAO

DEPT. OF ELECTRICAL ENGINEERING

ANDHRA UNIVERSITY,VISAKHAPATNAM,INDIA

ABSTRACT

A scheme for obtaining a linear feedback law for a linear discrete system as a result of minimising a quadratic performance index is considered in this paper. The resulting closedloop system has the property that all its poles lie inside a circle having a radius $e^{-\alpha}$ with centre at the origin of the Z - plane and $\alpha > 0$ may be chosen by the designer. Simulation studies are made to verify the method.

INTRODUCTION

In many applications, mere stability of a closedloop system is not enough. It is required that the poles of the closedloop system should lie in a certain restricted region of stability. Generally two methods are employed to derive a feedback control law for linear systems. The first approach, known as the pole placement technique [1,2], selects control law such that the closedloop system achieves certain desired locations for its poles. The second approach, based on optimal control theory [3] determines a control law which minimises a given performance index. A procedure for the selection of an optimal control law that minimises a given performance index and at the same time ensures that the closedloop poles lie in a prescribed region has been described by Anderson and Moore[4], Kawaski and Shimemura[5].

However, the continuous time results cannot be directly extended to the discrete time case. Solhiem [6] has developed a discrete version of Solhiem [7], but the optimality of the closedloop system is lost due to the difference in the form of the continuous and discrete riccati equations associated with the optimal regulator problem. Amin [8] took the same problem and the same recursive approach and derived an improved resut in which the optimality of the closedloop system is assured.

This paper presents a scheme for obtaining a linear feedback law for a linear discrete system as a result of minimising a quadratic performance index. The resulting closedloop system has the property that all its poles lie inside a circle having radius $e^{-\alpha}$ with centre at the origion of the Z-plane and $\alpha>0$ may be chosen by the designer.

PROBLEM FORMULATION AND SOLUTION

Consider the controllable linear time invariant discrete system:

$$x(k+1) = Px(k) + Gu(k) \qquad \ldots (1)$$

where x(k) and u(k) are n and m dimensional state and control vectors respectively. P and G are constant matrices of appropriate dimensions. The system (1) may be an originally discrete system or it may be a sampled data version of a continuous time system

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad \ldots (2)$$

In such a case the matrices P and G of (1) are given by

$$P = e^{(AT)} \qquad \ldots (3)$$

$$G = \int_0^T e^{[A(T-\tau)]}Bd\tau \qquad \ldots (4)$$

for a given discretization interval T.

Now let the system (1) is controlled in accordance with a linear feedback law

$$u(k) = -Fx(k) \qquad \ldots (5)$$

The dynamics of the controlled system is governed by

$$x(k+1) = (P-GF)x(k) \qquad \ldots (6)$$

There are two main approaches for determining the feedback matrix F, namely optimal control and modal control. The optimal control approach is based on the minimisation of the quadratic performance index

$$PI = \sum_{k=0}^{\infty} [x^T(k)Qx(k) + u^T(k)Ru(k)] \qquad \ldots (7)$$

where Q and R are non-negative and positive definite symmetric matrices respectively. The supersrcipt 'T' here denotes tranpose. The optimal feedback matrix is given by

$$F = R^{-1}G^TP^{-1}[S - Q] \qquad \ldots (8)$$

where S is the positive definite solution of the riccati equation

$$S - P^TSP + P^TSG[G^TSG + R]^{-1}G^TSP - Q = 0 \qquad \ldots (9)$$

The feedback matrix is unique for the given weighting matices Q and R. Since the system (1) is controllable, it is well known that the system (6) is stable, when optimally controlled, that is, the eigenvalues of (P-GF) are located inside the unit circle. One does not, however, know the exact location of the eigenvalues beforehand, and it may therefore happen that the controlled system although minimising the optimality criterion, does not have a desired degree of stability.

The second main approach for determining a feedback matrix F, namely modal control, assures acceptable stability through apriori specified eigenvalues of the controlled system (6). The feedback matrix F in this case is, however, not normally unique, and one cannot determine performance criteria (Q and R matrices) corresponding to particular modal feedback matrix. A particular modal design may infact be very wasteful in cotrol effort or cost, that is, controls may work against one another, while satisfying, of course, the eigenvalue specifications This is the most serious drawback of this approach.

In this paper we combine the best features of both the methods and design a feedback matrix which is optimal and also ensures prescribed degree of stability.

Consider the following perfomance index in place of (7),

$$PI_{\alpha} = \sum_{k=0}^{\infty} e^{2\alpha k}[x^T(k)Qx(k) + u^T(k)Ru(k)] \qquad \ldots (10)$$

which leads to a linear control law of the type (5), with the additional property that the closedloop system poles lie inside a circle with radius $e^{-\alpha}$ with centre at the origion of the Z-plane.

Defining

$$\hat{x}(k) = e^{\alpha k}x(k) \ ; \ \hat{u}(k) = e^{\alpha k}u(k) \qquad \ldots (11)$$

we obtain

$$\begin{aligned} \hat{x}(k+1) &= e^{\alpha(k+1)}x(k+1) \\ &= e^{\alpha(k+1)}[Px(k) + Gu(k)] \\ &= [\bar{P}\hat{x}(k) + \bar{G}\hat{u}(k)] \end{aligned} \qquad \ldots (12)$$

where $\bar{P} = Pe^{\alpha}$ and $\bar{G} = Ge^{x}$.

Using (11) the performance index (10) can be written as

$$PI_{\alpha} = \sum_{k=0}^{\infty}[\hat{x}^T(k)Q\hat{x}(k) + \hat{u}^T(k)R\hat{u}(k)] \qquad \ldots (13)$$

Now considering the system (12) and PI (13), construction of the control law is thus similar to a case where the degree of stability is not prescribed. The optimal control law that minimises (13) subject to the constraint (12) is given by

$$\hat{u}(k) = -F_{\alpha}\hat{x}(k) \qquad \ldots (14)$$

where

$$F_{\alpha} = R^{-1}G^T\bar{P}^{-T}[S_{\alpha} - Q] \qquad \ldots (15)$$

where S_{α} is the positive definite solution of the riccati equation

$$S_{\alpha} - \bar{P}^T S_{\alpha}\bar{P} + \bar{P}^T S_{\alpha}G[G^T S_{\alpha}G + R]^{-1}G^T S_{\alpha}\bar{P} - Q = 0 \qquad \ldots (16)$$

Using (11) and (14) it can be shown

$$u(k) = -F_{\alpha}x(k) \qquad \ldots (17)$$

Using (17) the closedloop form of system (1) becomes

$$x(k+1) = (P - GF_{\alpha})x(k) \qquad \ldots (18)$$

while the same for system (12) is

$$\hat{x}(k+1) = e^{\alpha}(P - GF_{\alpha})\hat{x}(k) \qquad \ldots (19)$$

Since all the poles of the closedloop system (19) are given by the eigenvalues of $e^{\alpha}(P-GF_{\alpha})$ and have the magnitudes less than unity. Hence it follows that the poles of the closedloop system (18) given by the eigenvalues of $(P-GF_{\alpha})$, lie with in a circle of radius $e^{-\alpha}$.

SIMULATION RESULTS

EXAMPLE-1

Consider the following 1st order example:

$P=0.9512$, $G=14.64$, $Q=1$, $R=1000$

$\alpha = 0$, $e^{-\alpha} = 1$, C.L. Pole = 0.6233

$\alpha = 0.693147$, $e^{-\alpha} = 0.5$, C.L. Pole = 0.2002

$\alpha = 1$, $e^{-\alpha} = 0.3679$, C.L. Pole = 0.1109

$\alpha = 10$, $e^{-\alpha} = 4.54 \times 10^{-5}$, C.L.Pole = 1.2×10^{-6}

EXAMPLE-2

Design of discrete-time linear regulator for seventh order synchronous machine is carried out using the results obtained in previous section. The closedloop poles for different values of α are shown in the table-1.

TABLE-1 LOCATION OF THE CLOSEDLOOP POLES

	Eigenvalues for different values of α					
$e^{-\alpha}$	λ_1	λ_2	λ_3, λ_4	λ_5	λ_6	λ_7
0.8333	-0.6281	0.6354	0.1038±j0.0124	-0.0328	0.0	0.0005
0.7143	-0.4623	0.4826	0.1088±j0.0124	-0.0328	0.0	0.0005
0.5016	-0.2271	0.2427	0.1038±j0.0124	-0.0328	0.0	0.0005
0.1999	-0.0370	0.0395	0.1038±j0.0124	-0.0321	0.0	0.0005
0.1003	-0.0327	0.0099	0.0953±j0.0105	-0.0091	0.0	0.0005

CONCLUSIONS

An approach to the design of discrete-time linear regulator with a prescribed degree of stability is presented. The analytical results obtained are verified through two examples. It is found that the prescribed degree of stability obtained using the proposed controller is quite satisfactory.

REFERENCES

1. Wonham, W.M., 1967, IEEE Trans. Atom. Control, 12, pp. 660.
2. Kailath, T., 1980, Linear Systems, Prentice-Hall,N.J.
3. Kalman, R.E., 1964, Am. Soc. Mech. Enggs., 86, pp. 1.
4. Anderson, B.D.O. and Moore, J.B., 1969, Proc. IEE, 116, pp. 2083.
5. Kawsaki, N. and Shimemura, E., 1983, Automatica, 19, pp. 557.
6. Solheim, O.A., 1974, Int. J. Control, 19, pp. 4117.
7. Solheim, O.A., 1972, Int. J. Control, 15, pp. 143.
8. Amin, M.H., 1984, Int. J. Control, 40, pp. 783.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 193-205

DECENTRALIZED FAULT DETECTION IN LARGE-SCALE SYSTEMS USING SECONDARY OBSERVERS

K.A. Gopala Rao
Department of Electrical ENGINEERING
A.U. Engineering College
Visakhapatnam 530 003

ABSTRACT

Due to the informational constraints, the computation and implementation of the Decentralized Fault Detection and Diagnosis (DFDD) algorithms are very complex. In this paper, we propose a new method for Decentralised fault detection(DFD) in large linear dynamical systems using secondary observers scheme.

1. INTRODUCTION

Failure (fault) detection and diagnosis (FDD) is currently the subject of extensive research and is an integral part of the design of highly reliable control systems [1-3]. Of the several methods available in literature for FDD, the observer based methods are gaining importance because the states generated can be used not only for FDD but also for feedback controller. Under normal conditions the state estimates can be used directly or else some compensation is provided under fault conditions before feeding them to the controller. However, most of the available FDD methods are based on the assumption that the information acquisition is centralised. This assumption does not hold in most practical situations, since for reasons of economy and reliability the information is spatially distributed. Thus, in many large systems the information available to fault detectors is localised and each detector monitors a specified section of the large system. Thus, decentralised fault detection methods gain importance. However, in limited information situations, like the above, application of the methods available for centralised systems is not straight forward and require new approaches. There is very little literature on this topic [4,5,6].

In this paper, the problem of designing Decentralised Failure Detection Systems (DFDS) for the detection of local faults in large linear dynamical systems using observers is addressed. We assume that fault detection and diagnosis are performed by local detectors (primary and secondary observers) - one associated with each subsystem.Further it is assumed that the local detectors use only the local information and donot communicate among themselves. Thus each subsystem considers the interactions of other systems as unknown inputs and the primary observers are designed based on the unknown input observer theory. Because of this some information regarding global faults at local station will be lost. Moreover in the observer based FDD the n-dimensional observer state error vector carries the necessary information on all possible failures. However, the problem is that in practical systems, in which only some of the components of the state vector are measured, one hase access only to the m-dimensional observer-output error vector, with $m<n$. In order to cope with the above two cases a secondary observer is designed to reconstruct the entire observer state error vector from the observer output error vector.

2. PROBLEM FORMULATION

Consider the large interconnected system comprising of N-subsystems described by,

$$x_i(k+1) = A_i x_i(k)+B_i u_i(k)+\sum_{j=1}^{N} A_{ij} x_j(k); i = 1,2,..,N \qquad (2.1)$$

$$y_i(k) = C_i x_i(k); \ i = 1,2,...,N \qquad (2.2)$$

$$y_{i0}(k) = C_{i0}\ x_i(k); \quad i = 1,2,...,N \qquad (2.3)$$

where $x_i(k)$ is an $n_i x1$ state vector, $u_i(k)$ is an $m_i x1$ input vector, $y_i(k)$ is a $p_i x1$ output vector of the ith subsystem. y_{i0} is a p_{i0} output vector of the ith subsystem.

Here we are concerned with the problem of designing N-local detectors for detecting abrupt changes in the N-local systems. The local detectors do not communicate amongst themselves. We have, available at each local detector, the following:

1. measurements of the local input $u_i(k)$ and local outputs $y_i(k)$ and $y_{i0}(k)$,
2. the nominal values of the coefficient matrices C_i, A_i, B_i and A_{ij}, and

In the equation (2.1), if $A_{ij} = 0$ or information about all x_j's is available at the ith detector, then the problem is a classical one. Otherwise, each interaction term $A_{ij}x_j$ acts as an unknown input which makes the detection problem difficult to solve. By appropriately redefining various quantities, we can rewrite the term $\Sigma A_{ij}x_j(k)$ as $D_i d_i(k)$ with D_i having full column rank q_i in all the subsystems. We note that the construction of a local detector to detect local faults in (2.1) - (2.3) requires the solution of an unknown input detector problem. The (2.1) is modified as

$$x_i(k+1) = A_i x_i(k)+B_i u_i(k)+D_i d_i(k); i=1,2,..,N \qquad (2.4)$$

We consider this model for further analysis in the following sections.

3. STRUCTURE OF THE PRIMERY OBSERVER

Consider the system (2.4), (2.2) and (2.3) . We assume that (C_i, A_i) is an observable pair,rank $C_i = p_i$, rank $D_i = q_i$ and rank $C_iD_i = q_i$.

For convenience we drop the subscript i in all the variables indicating the ith subsystem and wherever necessary we explicitly include the subscript i.

Since C has full row rank,a suitable coordinate transformation on the states can be found such that C has the representation $[I_p,0]$, I_p being the pxp unit matrix. So without loss of generality, we assume that the given C is in this form.

Now an (n-p)-th order observer can be constructed as follows[7]:

$$Z(k+1) = FZ(k) + Hy(k) + Gu(k) \tag{3.1}$$

where

$$Z(k) = Tx(k) + \epsilon(k) \tag{3.2}$$

with

$$FT + HC = TA \tag{3.3}$$

$$TB = G \tag{3.4}$$

$$TD = 0 \tag{3.5}$$

If (3.1) is to be a minimal order observer, we must also have $[C^T,T^T]^T$ non singular. The particular structure of C that we have chosen logically allows one to choose T as

$$T = [-L \quad I_{n-p}] \tag{3.6}$$

where L is an (n-p) x p arbitrary matrix which will be chosen to satisfy certain conditions stated later.

Let

$$\begin{bmatrix} C \\ T \end{bmatrix} = [V \quad P] \tag{3.7}$$

where

$$V = \begin{bmatrix} I_p \\ L \end{bmatrix} \quad ; \quad P = \begin{bmatrix} 0 \\ I_{n-p} \end{bmatrix} \tag{3.8}$$

We partition A,B,D conformably with C,T,V and P as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad ; \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \quad ; \quad D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$$

Then we have the following explicit expressions for F,H and for the conditions (3.3) - (3.5) [8,7]

$$F = A_{22} - LA_{12} \tag{3.10}$$

$$H = FL + A_{21} - LA_{11} \tag{3.11}$$

$$G = B_2 - LB_1 \tag{3.12}$$

$$D_2 - LD_1 = 0 \tag{3.13}$$

Observer (3.1) can be implemented, if[7]

1. $D_2 - LD_1 = 0$ (3.14)

   and

2. $(A_{22} - LA_{12})$ has stable eigenvalues (3.15)

There exists an L such that (3.14) - (3.15) hold, if rank CD = rank D = q and $p \geq q$[7].

It has been shown in Kudva et al [7] that the eigen spectrum of $(A_{22} - LA_{12})$ consists of two parts:

i) a part which is fixed at the invariant zeros of the triple (C,A,D)

ii) a freely assignable part.

Thus an unknown input observer exists if the invariant zeros of (C,A,D) are stable and $p \geq q$[7].

So in our problem we further assum that the invariant zeros of the triple (C,A,D) are stable.

Now from (2.2), (3.2) and (3.7) we get

$$\hat{x}(k) = PZ(k) + VY(k) \tag{3.16}$$

Defining

$$e(k) = x(k) - \hat{x}(k) \tag{3.17}$$

Then we can show fr.m (3.2) and (3.16)

$$e(k) = -P\epsilon(k) \tag{3.18}$$

We can further show that:

$$\epsilon(k+1) = -TAe(k) \tag{3.19}$$

Thus we get

$$e(k+1) = PTA\ e(k) \tag{3.20}$$

Under normal operation $e(k) \rightarrow 0$ after an initial transient and when a fault occurs it is nonzero and thus indicates the fault. In the following sections we consider the fault models in which we are interested and how the information of these faults propagate into the estimation error.

4. FAULT MODELS

The faults we consider here are abrupt changes of unknown magnitudes occuring in local systems at unknown times. The changes in the local system can occur in a number of ways. These can be failures of sensors, actuators or large plant parameter drifts. The failure detection and diagnosis considered here is based on modelling of failures by including in the nominal model (2.1) - (2.3), additive impulse or step disturbances, occurring with unknown magnitudes at unknown times. More specifically, the following four different types of failure modes can be considered:

Type 1: State jump

$$x_i(k+1) = A_i x_i(k) + B_i u_i(k) + \sum_{j=1}^{N} A_{ij} x_j(k) + \tau_i \, \delta_{k+1,\theta} \qquad (4.1)$$

Type 2: State step

$$x_i(k+1) = A_i x_i(k) + B_i u_i(k) + \sum_{j=1}^{N} A_{ij} x_j(k) + \tau_i \, \sigma_{k+1,\theta} \qquad (4.2)$$

Type 3: Sensor jump

$$y_i(k) = C_i x_i(k) + \tau_i \, \delta_{k,\theta} \qquad (4.3)$$

Type 4: Sensor step

$$y_i(k) = C_i x_i(k) + \tau_i \, \sigma_{k,\theta} \qquad (4.4)$$

In all these models, θ is the unknown time of occurrence of the failure assuming a finite value when the failure occurs and $+\infty$ if there is no failure. τ_i is the unknown failure vector which is deterministic. $\sigma_{k,\theta}$ and $\delta_{k,\theta}$ are unit step and Kronekar delta functions respectively, i.e.

$$\sigma_{k,\theta} = \begin{cases} 0 & k < \theta \\ 1 & k \geq \theta \end{cases}$$

$$\delta_{k,\theta} = \begin{cases} 1 & k = \theta \\ 0 & \text{otherwise} \end{cases}$$

5. FAULT DETECTION USING SECONDARY OBSERVER

In order to observe how these fault information evolves in the estimation error, we consider type 1 fault model, i.e. (4.1), in this section. For other fault models similar expressions can be derived.

If we implement the observer designed in section 3 with the system model (4.1), i.e. under fault conditions we can show that e_f, the estimation error under fault conditions will contain two components (i) the estimation error e under normal conditions and (ii) a fault component, i.e. if the fault occurs at $j = \theta$ then for $j \geq \theta$

$$e_f(j) = e(j) + PT(APT)^{j-\theta} \tau \tag{5.1}$$

and when $j < \theta$ the term due to fault is zero.

However, neither e nor e_f is available for measurement. We use the measurement equation (2.3) and an estimate of y_0 to generate the required residuals as shown.

$$\tilde{e}_f(j) = y_0(j) - \hat{y}_0(j) = C_0 \ e(j) + C_0PT(APT)^{j-\theta} \tau \tag{5.2}$$

which can be written as

$$\tilde{e}_f(j) = \tilde{e}(j) + \tilde{e}_{fault}(j) \tag{5.3}$$

It is evident from (5.2) that under normal conditions the $\tilde{e}_f$ is zero and is equal to $C_0PT(APT)^{j-\theta}\tau$ under fault condition. Thus we can use this information to identify the fault.

Here we present an approach which uses the two equations (5.1) and (5.2). It is easy to show that these can be generated from the following dynamic system under fault conditions:

$$e(k+1) = \bar{A}e(k) + \bar{\tau}_{k+1,\theta} \tag{5.4}$$

$$\tilde{e}(k) = C_0e(k) \tag{5.5}$$

where

$$\bar{A} = PTA \tag{5.6}$$

$$\bar{\tau} = PT\tau \tag{5.7}$$

An observer will be designed based on the system (5.4) and (5.5) with the term $\bar{\tau}$ omited to estimate e(k). This is known as secondary observer. We propose for the secondary observer the following structure:

$$\hat{e}(k+1) = \bar{A}\ \hat{e}(k) + K\ [\tilde{e}(k) - C_0\ \hat{e}(k)] \tag{5.8}$$

Let us know define the secondary observer's error vector $\alpha(k)$, as follows:

$$\alpha(k) = e(k) - \hat{e}(k) \tag{5.9}$$

Using (5.1), (5.2), (5.8) and (5.9) we can show

$$\alpha(k+1) = [\bar{A} - KC_0]\ \alpha(k) \tag{5.10}$$

If the eigenvalues of the nxn matrix $(\bar{A}-KC_0)$ are adequately located in the unit circle, the solution of (5.10) will be asymptotically stable and will vanish as $k \to \infty$ under unfaulted condition. To satisfy this condition the nxp_0 gain matrix K is chosen properly.

The output of the secondary observer $\hat{e}(k)$, which is the reconstructed error vector of the state, will follow e(k) after a initial transient under normal condition. However under fault condition, since the secondary observer did not take into account the input $\bar{\tau}$ of (5.4), the estimate will not follow the true error. This difference in e(k) and $\hat{e}(k)$ will be used to indicate the fault condition.

6. EXAMPLE

Consider the following large-scale interconnected system with subsystems described by the following matrices:

Subsystem 1:

$$A_{11} = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0.7 & 0 & 0 & 0 \end{vmatrix} ; \Sigma A_{1j} X_j = \begin{vmatrix} 0 & 0 & -0.5 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \begin{vmatrix} x_5 \\ x_9 \\ x_{10} \end{vmatrix}$$

$$C_1 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

The local states are x_1, x_2, x_3 and x_4 and local outputs are y_1, y_2 and y_3.

Subsystem 2:

$$A_{22} = \begin{vmatrix} 0 & 0 & 0 & 0.8 \\ 0 & 0 & 0 & -0.6 \\ 0 & 0 & 0.5 & 5 \\ 0 & 0 & 0 & 0 \end{vmatrix} ; \Sigma A_{2j} x_j = \begin{vmatrix} 0 \\ 9 \\ 0 \\ 10 \end{vmatrix} x_{10}$$

$$C_2 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

The local states are x_5, x_6, x_7 and x_8 and outputs are y_4, y_5 and y_6.

Subsystem 3:

$$A_{33} = \begin{vmatrix} 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.6 \\ 0 & 0 & -0.2 & 0 \end{vmatrix} ; \Sigma A_{3j} X_j = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 10 & 0 & 0.8 & 0 \\ 0 & 0.4 & 0 & 5 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ x_5 \\ x_6 \end{vmatrix}$$

$C_3 = I_4$. The local states are x_9, x_{10}, x_{11} and x_{12} and outputs are y_7, y_8, y_9 and y_{10}.

For the above systems, the matrices D_i and the vectors $d_i(k)$ of equation (2.4) become

$$D_1 = \begin{vmatrix} 0 & -0.5 \\ 0.70711 & 0 \\ 0 & 0 \\ 0 & 0 \end{vmatrix} \quad ; \quad d_1(k) = \begin{vmatrix} \dfrac{x_5(k) + x_9(k)}{x_{10}^2(k)} \end{vmatrix}$$

$$D_2 = \begin{vmatrix} 0 \\ 9 \\ 0 \\ 10 \end{vmatrix} \quad ; \quad d_2(k) = x_{10}(k)$$

$$D_3 = \begin{vmatrix} 0 & 0 & 0 \\ 0.5 & 0 & 0 \\ 10 & 0 & 0.8 \\ 0 & 0.0319 & 0 \end{vmatrix} ; d_3(k) = \begin{vmatrix} x_1(k) \\ 0.079745x_2(k) + 0.99682x_6(k) \\ x_5(k) \end{vmatrix}$$

For the purpose of demonstration of DFD, we concentrate on subsystem 2. The additional measurement equation (2.3) for this case is

$$y_{20}(k) = \begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} x_5(k) \\ x_6(k) \\ x_7(k) \\ x_8(k) \end{vmatrix}$$

We considered a jump type of fault (type 1) of magnitude 10 in the state variable x_7 of the subsystem 2 occurring at $\theta=5$ and observed that the detector of subsystem designed using the technique of sections 3 and 4 give the following results:

A change of -10.02 is observed at $\theta = 5$ in $\hat{e}(k)$. When the system is normal the detector gave the following result:

$\hat{e}(k) = 0$ after an initial transient.

To show that the detector of subsystem 2 will not detect faults in other subsystems we created a jump type of fault in state variable x_{10} of subsystem 3 of magnitude 100 at $\theta = 5$. The observations made at detector of subsystem 2 indicate the following:

$\hat{e}(k)$ is observed to be the same as under normal conditions.

CONCLUSIONS

In this paper, we have developed a new technique for local fault detection in a decentralized environment, with the informational constraints stated in section 2, using unknown input observer theory. The combination of primary and secondary observers is optional, i.e., one can use simple techniques on equation (5.2) for DFD. However, the use of secondary observer estimates the state estimation error e(k), which contains lot of information regarding the system faults. Although we have considered specific fault models, our method is general and can also be applied for other types of faults [9,10].

REFERENCES

1. A.S. Willsky, "A survey of design methods for failure detection in dynamic systems", Automatica, Vol. 12, No. 6, Nov. 1976, pp. 601-611.

2. R.N. Clark, 'The dedicated observer approach to instrument failure detection", Proc. 18th IEEE Conf. on Decision and Control, 1979, pp. 237-241.

3. R. Isermann, "Process fault detection based on modeling and estimation methods - A survey", Automatica, Vol. 20, No. 4, 1984, pp. 387-404.

4. M.G. Singh, M.F. Hassan, Y.L. Chen, D.S. Li and Q.R. Pan, "New approach to failure detection in large systems", IEE Proc., Vol. 130, Pt. D, No. 5, 1983, pp. 243-249.

5. M. Hodzic and D.D. Siljak, "Failure detection in large sparse systems", Large-Scale Systems, Vol. 9, 1985, pp 83-100.

6. K.A. Gopala Rao, "Fault detection and diagnosis in linear systems", Ph.D. thesis, 1986, I.I.Sc., Bangalore.

7. P. Kudva, N. Viswanadham and A. Ramakrishna, "Observers for linear systems with unknown inputs", IEEE Trans. on Automatic Control, AC-25, No. 1, 1980, pp. 113-115.

8. B. Gopinath, "On the control of linear multiple input - output systems", Bell System Tech. J., Vol.50, No.1, 1971, pp. 1063-1081.

9. R.V. Beard, "Failure accommodation in linear systems through self reorganization", Ph.D. Thesis, M.I.T., Cambridge, Mass., Feb. 1971.

10. H.L. Jones, "Failure detection in linear systems", Ph.D. Thesis, M.I.T., Cambridge, Mass., 1973.

*Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 207-218*

APPLICATION OF PATTERN RECOGNITION TO POWER SYSTEMS FOR TRANSIENT SECURITY EVALUATION

DR. C. M. ARORA
Dept. of Electrical Engg.,
M R Engineering College,
Jaipur-302017, INDIA.

DR. S. L. SURANA
Dept. of Electrical Engg.,
M B M Engineering College,
Jodhpur-342011, INDIA.

Abstract - For effective control of power systems a quick transient security evaluation of the operating state is essential. The paper presents a pattern recognition (PR) method for quick transient security evaluation of power systems using steady state variables. The method uses a pattern vector which is selected on the basis of power system governing equations. The presented approach requires very small on-line computational time while keeping misclassification rate within desired limits. The method is illustrated by an example.

1 Introduction

The primary aim of power system control is to provide adequate uninterrupted (reliable) supply of electrical power within set limits of frequency and voltage levels. To ensure reliable operation of modern power systems, which are large and complex, a quick evaluation of power system security is essential.[1-3]

The transient security of a power system is a dynamic problem[2,4]. In the decade 1970-80 different pattern recognition (PR) techniques have been studied for transient security evaluation using steady state variables which resulted in classification error typically of the order of 10%[2,4]. In the decade 1980-90 various PR techniques have been studied for transient security evaluation using transient variables[4-8]. This approach though reduced the classification error considerably, resulted in complex methods requiring more computational time.

In this paper it is pointed out that the large classification error in transient security evaluation using steady state variables as reported in the literature is due to the application of inappropriate/inadequate PR methods. A pattern vector containing minimum number of independent and directly measurable steady state variables for transient security evaluation is derived. A PR method is presented using this pattern vector which is simple, accurate enough and requires smaller on-line and off-line computational time and efforts as compared to the existing methods. Application of the proposed method is illustrated by an example.

2 Proposed pattern recognition method

The presented approach requires designing of separate security classifiers for different network and generator configurations. It will be seen later that this approach drastically reduces the size of the classifier. This approach also makes parallel processing of all the configurations possible thereby enabling fast determination of security classifiers.

The proposed procedure for determining a transient security classifier using PR method requires defining of a pattern vector $\underline{x}$ whose components contain all significant variables of the power system. The vector alongwith the power system transient security are evaluated at many representative operating conditions to generate a set of known class (security) patterns, called training set. The transient security is evaluated by performing load flow calculations using economic dispatch and transient stability studies.

Some components of the pattern vector may be strongly correlated with one another and hence all the components may not be necessary to obtain the security classifier. Therefore a process of dimensionality reduction is performed to identify significant and hopefully uncorrelated set of components. The set of significant and uncorrelated components is called feature subset $\underline{z}$ and the process of obtaining it is called feature selection. Final step is to determine a security classifier function $S(\underline{z})$ by classifier design algorithm such that $S(\underline{z}) \geq 0$ for a secure state and $S(\underline{z}) < 0$ for an insecure state.

The proposed pattern vector, feature selection algorithm and classifier design algorithm are discussed in subsequent sections.

Pattern vector

Power system variables which are closely related to transient security are selected as components of pattern vector.[4-7] Generally, for the determination of the power system security first swing stability of the multimachine power system for each fault in contingency set is evaluated. The criterion for the evaluation of first swing stability is that at the end of one second of occurrence of the disturbance if the generators' rotor angles differ less than or equal to 180° then the system is termed as stable and if the generators' rotor angles differ more than 180° then the system is termed as unstable. If the system is stable for all the disturbances in the contingency set then it is termed as secure. If the system is unstable for one or more disturbances in the contingency set then it is termed as insecure.[2,4,5]

Thus the transient security of power system concerns with the generators' rotor angles deviations. The generators' rotor angles are determined from the following set of swing equations[4-7].

$$M_i \frac{d^2\delta_i}{dt^2} = P_{mi} - P_{ei} \quad , \ i=1,2,..,m \tag{1}$$

$$P_{ei} = E_i^2 Y_{ii} \operatorname{Cos} \Theta_{ii} + E_i \sum_{\substack{j=1 \\ j\neq i}}^{m} E_j Y_{ij} \operatorname{Cos}(\delta_i-\delta_j-\Theta_{ij}) \tag{2}$$

where,

M_i = coefficient of inertia of i th generator;

δ_i = rotor angle of i th generator assumed equal to angle of voltage behind transient reactance;

P_{mi} = mechanical power input to i th generator;

P_{ei} = electrical power output of i th generator;

E_i = voltage magnitude behind transient reactance X_d' of i th generator;

Y_{ij} = magnitude or modulus of ij th element of equivalent system admittance matrix (system represented by generator buses only);

Θ_{ij} = angle or argument of ij th element of equivalent system admittance matrix; and

m = number of generators.

The Eqs. (1) and (2) indicate that the generators' rotor angles and hence transient security is a function of the following variables:

(i) Inertia constants of the generators;

(ii) Predisturbance generators' rotor angles which are functions of the predisturbance terminal voltages of the generators, their transient reactances and predisturbance real and reactive power outputs of the generators;

(iii) Mechanical power input to each generator which is assumed to be equal to predisturbance electrical output (real power) of each generator;

(iv) Magnitude of internal voltage (voltage behind transient reactance) of each generator which is a function of its predisturbance terminal voltage, transient reactance and predisturbance real and reactive power output;

(v) Modulus and argument of the power system admittances which are a function of impedances and susceptances of transmission lines, transformers and static capacitors during disturbance and post-disturbance periods, loads and predisturbance voltage magnitudes of the buses to which the loads are connected; and

(vi) Disturbance duration and maximum time of study (maximum time of study is generally taken as one second for first swing stability).

Therefore, the features which can be considered for transient security classifier are inertia of each generator, predisturbance terminal voltage (complex) of each generator, transient reactance of each generator, predisturbance real and reactive power outputs of each generator, impedances and susceptances of transmission lines, transformers and static capacitors during disturbance and post-disturbance periods, real and reactive loads and predisturbance voltage magnitude of each load bus, disturbance duration and maximum time of study.

As mentioned in the literature[9] the features which are common to all pattern classes under consideration carry no discriminatory information and hence can be ignored. Such features are inertia constant of each generator, transient reactance of each generator, impedances and susceptances of transmission lines, transformers and static capacitors during disturbance and post-disturbance periods, duration of the disturbance and maximum time of study. These features are common to both secure class and insecure class for the given power system and contingency set and hence contain no discriminatory information and can be ignored. The remaining features which are important from classification point of view are :

(i) Predisturbance terminal voltage (complex) of each generator,
(ii) Predisturbance real and reactive power output (generation) of each generator,
(iii) Real and reactive loads on each bus,
(iv) Predisturbance voltage magnitude of each load bus.

It may be noted that all of these remaining features are steady state variables only. Hence it can be concluded that it is sufficient to use steady state variables as features for transient security evaluation.

In a power system operating in steady state the terminal voltage (complex) of generators, reactive power output of generators and voltage magnitude of each load bus are functions of real and reactive power demand/load at each bus, voltage magnitude of generator buses, reactive power limits of generators and network admittances.[10,11] Therefore the set of variables obtained for pattern vector may be modified as follows :

(i) Real power generation at each generator bus, PG_i,

$i=1,2,\ldots,m$.

(ii) Real power demand at each bus, PL_i, $i=1,2,\ldots,n$.

(iii) Reactive power demand at each bus, QL_i, $i=1,2,\ldots,n$.

where n is the total number of buses in the system.

Voltage magnitude of generators, reactive power limits and network admittances (complex) have not been included in the above set because they are common to secure and insecure classes for the given power system and hence do not provide discriminatory information from classification point of view.

When the system is economically dispatched the real power generation at each generator bus is a function of real power demand/load at each bus[10]. Also it may be assumed that the power factor at system load buses remain practically constant[4,12] thereby making reactive power demand of each load bus a function of real power demand. From these considerations it is clear that the pattern vector containing real power demand at each bus PL_i, $i=1,2,\ldots n$ is sufficient to describe operating state of a given economically dispatched power system for transient security evaluation.

Assuming a quadratic classifier the pattern vector will consist of the real power demand at each bus PL_i, $i=1,2,\ldots,n$ and their products $PL_i \times PL_j$, $i=1,2,\ldots,n$, $j=i,i+1,\ldots,n$. The number of components in the pattern vector is $n(n+3)/2$. Expressed mathematically the pattern vector chosen is

$$[PL_1\ PL_2\ ..\ PL_n\ PL_1 \times PL_1\ ..\ PL_1 \times PL_n\ ..\ PL_n \times PL_n\ ]^T$$

Each of the component of the pattern vector is divided by its average value in the training set before applying feature selection algorithm. The transformation converts each variable to have unit average value in training set. It has been observed that the feature selection and classifier design algorithms work effectively by using this transformation. After classifier design the weights[1,2,9] of the classifier obtained are transformed by dividing into weights the average value of the corresponding variable in the training set. This makes the application of the classifier on test patterns (test load samples) without transforming them.

Feature selection

The proposed feature selection procedure is implemented in two stages. In the first stage the features are selected from pattern vector $\underline{x} = [x_1 \; x_2 \; .. \; x_q]^T$ based on maximization of a criterion function F defined by Eq. (3) :

$$F = \frac{| \; m_s - m_i \; |}{s_s + s_i} \tag{3}$$

where,

m_s = mean of the variable in the secure class

m_i = mean of the variable in the insecure class

s_s^2 = variance of the variable in the secure class

s_i^2 = variance of the variable in the insecure class

Selection of features begin with the computation of F-values for all the components (variables) of pattern vectors in the training set. The variable with largest F-value is selected as first feature, let this variable be designated as z_1. The correlation coefficients between z_1 and each of the rest of the variables are determined. Those variables are discarded which are highly correlated to z_1 say with correlation coefficient greater than k_c where $0 \leq k_c \leq 1.0$, thereby removing most of the redundant information. Now from the remaining variables, the one with largest F- value is selected as second feature say z_2. The preceding procedure is repeated until all the variables are considered. This process will give a feature set $\underline{z} = [\; z_1 \; z_2 \; .. \; z_p \;]^T$ ordered in descending order of value of F-function of each feature. The selection of the value of correlation coefficient k_c is important and should be taken such as to give optimal results.

In the second stage a feature subset is selected from the features obtained in the first stage using performance dependent feature selection procedure. Starting from the highest ranked feature z_1 in the feature subset the remaining features $z_2, z_3, .., z_p$ are included in the feature subset one by one, in the order of their ranking, and an optimal classifier corresponding to each of the feature subset generated is designed using proposed classifier design algorithm. The classifier designed is tested on training set patterns everytime and the process is terminated when either error free classifier is obtained or inclusion of the feature increases classification error. The above process gives a feature subset $\underline{z} = [z_1\ z_2\ ..\ z_l]^T$, $l \leq p$.

Classifier design algorithm

Increment-correction algorithm[9] based on statistical approach is used for determination of security function or classifier. The algorithm provides an optimum classifier for the given feature subset.

3 Illustrative example

Power system description

A 9 bus, 230 kv power system[13] is chosen to illustrate the application of the proposed transient security evaluation method. Fig. 1 depicts a one-line diagram of this system. The system parameters in p.u. on 100 MVA base are given in Tables 1 to 3.

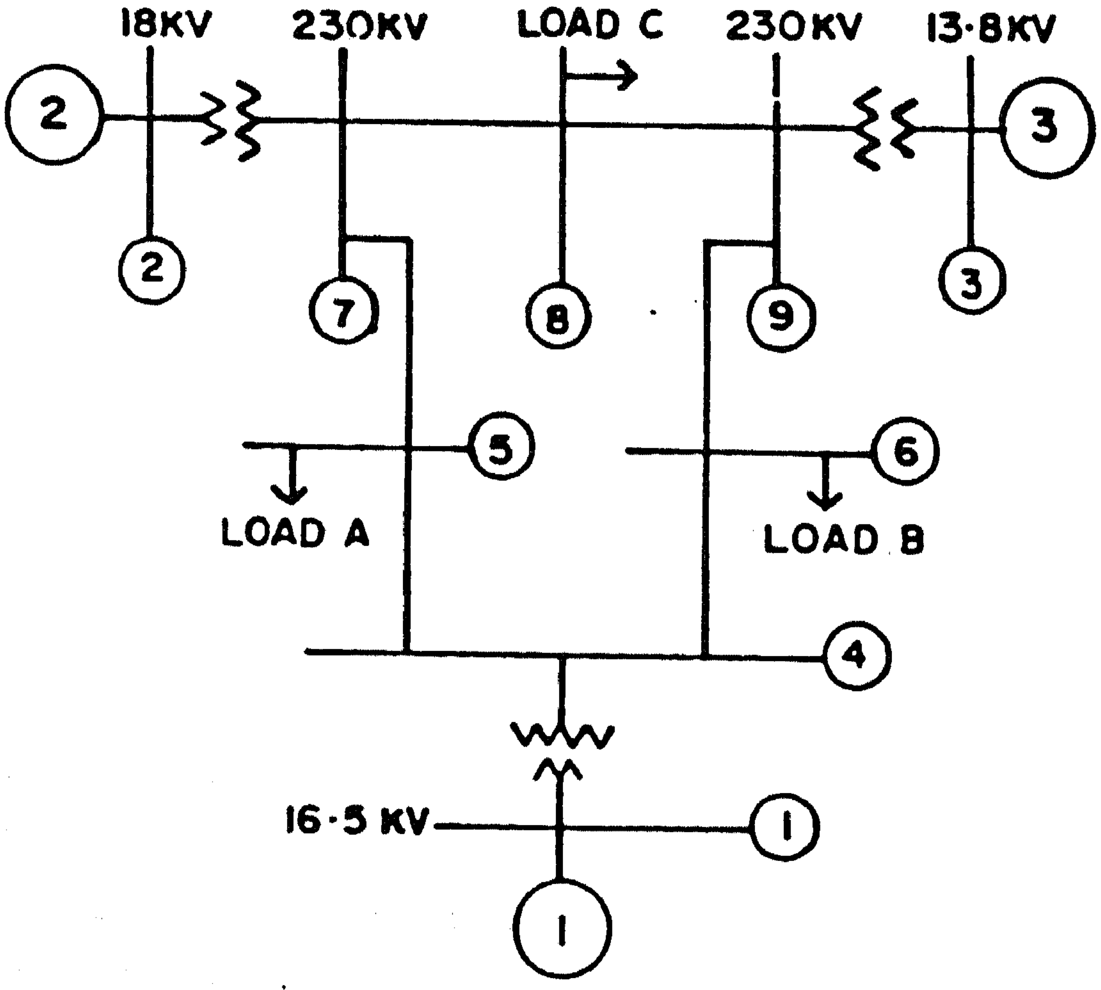


FIG. 1 ONE LINE DIAGRAM OF 9 BUS SYSTEM

Table - 1

Transformer Data (in p.u.)

From Bus	To Bus	Resistance	Reactance
2	7	0.0	0.0625
3	9	0.0	0.0586
1	4	0.0	0.0576

Table - 2

Transmission Line Data (in p.u.)

From Bus	To Bus	Resistance	Reactance	Susceptance
7	8	0.0085	0.0720	0.1490
8	9	0.0119	0.1008	0.2090
9	6	0.0390	0.1700	0.3580
6	4	0.0170	0.0920	0.1580
4	5	0.0100	0.0850	0.1760
5	7	0.0320	0.1610	0.3060

Table - 3

Generator Data

Generator	1	2	3
Rated MW	247.5	192.0	128.0
X_d' (p.u.)	0.0608	0.1198	0.1813
Stored-energy	2364	640	301
Cost function $ per hr	$59.4+0.0703P+0.00745P^2$	$155.5+0.489P+0.00393P^2$	$77.3+0.456P+0.00773P^2$

Note : 1. Here P is the generated power in MW

2. Stored energy is in MW-s at rated speed

It is assumed that the loads are randomly distributed and have a multivariate normal distribution function[14] with the following mean and covariance.

Mean load = $[\ PL_5\ PL_6\ PL_8\ ]^T = [\ 1.25\ 0.90\ 1.0\ ]^T$
(in p.u.)

$$\text{Covariance of load} = \begin{array}{c} \\ 5 \\ 6 \\ 8 \end{array} \begin{array}{c} \begin{array}{ccc} 5 & 6 & 8 \end{array} \\ \left[\begin{array}{ccc} 0.01563 & 0.00900 & 0.01000 \\ 0.00900 & 0.00810 & 0.00720 \\ 0.01000 & 0.00720 & 0.01000 \end{array}\right] \end{array}$$

The power factor of the loads are assumed to be constant and equal to 0.94. Bus 1 is taken as the swing bus with voltage magnitude 1.025 p.u. and the buses 2 and 3 are voltage controlled buses with voltage magnitude 1.025 p.u..

Type and duration of the disturbance : The contingency set contained only one disturbance that is a three phase short circuit on transmission line connecting buses 5 and 7, near bus 7. The fault is cleared after 200 milliseconds by opening of the line at both ends. On conducting detailed stability studies it is found that three phase faults at other locations are less severe.

Security Function Evaluation and Test Results

Ninety different load samples are generated for loads A, B and C according to given means and covariances of loads. The transient security class of each load sample (representing operating state of the system) is calculated as follows. Each load sample is considered one by one. For the load sample under consideration the optimal output of each generator is determined by economic dispatch[10] calculations including losses. Using optimal output of generators a load flow calculation is performed to obtain steady operating state of the system for the load sample. Gauss-Seidel iterative technique[11,15] is used for obtaining load flow solution. The transient stability of the system operating state under the specified disturbance is determined by solving swing equations using modified Euler's method[11]. If the operating state of the power system is found to be stable then the load sample is classified as secure. Otherwise, the load sample is classified as insecure.

A pattern vector is formulated by the loads and their products for each of the 90 load samples. Since there are three loads, the pattern vector contained nine components in all. Application of the feature selection algorithm gives two features namely $PL_5 \times PL_6$ and $PL_5 \times PL_8$. The limiting value of coefficients of correlation k_c during feature selection is taken equal to 0.95. Using the subset of features a classifier is designed by the classifier design algorithm. The classifier obtained is as follows.

$$S(\underline{PL}) = -0.2495536 \times PL_5 PL_6 - 0.2455548 \times PL_5 PL_8 + 0.7436354$$

The classification rule is :

if $S(\underline{PL}) > 0.0$ then the set $[\ PL_5\ PL_6\ PL_8\ ]^T$ is secure,
if $S(\underline{PL}) < 0.0$ then the set $[\ PL_5\ PL_6\ PL_8\ ]^T$ is insecure, and
if $S(\underline{PL}) = 0.0$ then the set $[\ PL_5\ PL_6\ PL_8\ ]^T$ is inconsistent.
Where all PL_i are in p.u..

The classifier is tested on all the 90 load samples and it is found that these are correctly classified giving a correct classification rate of 100%. The classifier is also tested on a separate set of 600 load samples constituting the test set and it is found that 589 load samples are correctly classified giving a correct classification rate of 98.2%.

4 Conclusions

In this paper use of steady state variables as features for the purpose of transient security evaluation of power system by PR method has been investigated. The presented PR method is simple, accurate enough and requires smaller on-line and off-line computational time and efforts as compared to existing methods. Therefore the proposed method can be used for quick on-line transient security evaluation of power systems enabling effective control of power systems.

5 References

[1] C. K. Pang, A. J. Koivo and A. H. El-Abiad, "Application of pattern recognition to steady state security evaluation in a power system" IEEE trans. on Systems, Man and Cybernetics,

Vol-SMC-3, pp 622-631, Nov. 1973.

[2] C. K. Pang, F. S. Prabhakara, A. H. El-Abiad and A. J. Koivo, "Security evaluation in power systems using pattern recognition" IEEE trans. on Power Apparatus Systems, Vol. PAS-93, pp 969-976, May/June 1974.

[3] A. K. Sinha and I. J. Nagarath, "Pattern recognition method for power system steady state security assessment" IE(I) Journal-EL-Vol. 64, pp. 269-273, April 1984.

[4] H. Hakimmashhadi and G. T. Haydt, "Fast transient security assessment" IEEE trans. on Power Apparatus Systems, Vol-PAS-102, no. 12, pp. 3816-3824,Dec. 1983.

[5] S. Yamashiro, "On-line secure-economy preventive control of power systems by pattern recognition" IEEE trans. on Power Systems, Vol-PWRS-1, no. 3, pp 214-219, August 1986.

[6] A. V. Machias, "A fuzzy pattern recognition approach to power system transient stability assessment" Electric Power Systems Research, Volume-16, pp. 141-146, 1989.

[7] A. V. Machias, "Transient stability evaluation by a pattern recognition approach using Multivalent Logics" Electric Power Systems Research, Volume-17, pp. 209-217, 1989.

[8] D. R. Ostojic and G. T. Haydt, "Transient stability assessment by pattern recognition in frequency domain" IEEE trans. on Power Systems, Vol. 6, No. 1, pp 231-237, February 1991.

[9] J. T. Tou and R. C. Gonzalez, "Pattern Recognition Principles" Addison-Wesley publishing company, 1974.

[10] A. R. Burgen, "Power System Analysis" Prentice-Hall, 1986.

[11] M. A. Pai, "Computer Techniques in Power System Analysis" Tata McGraw Hill publishing company Limited, 1979.

[12] C. S. Chang, T. S. Chung and K. L. Lo, "Application of pattern recognition technique to power system security analysis and optimization" IEEE trans. on Power Systems, Vol. 5, no. 3, pp 835-841, Aug. 1990.

[13] P. M. Anderson and A. A. Faud, "Power System Control and Stability" The Iowa state university press, Iowa, 1977.

[14] K. Miller, "Multidimensional Guassian Distributions" J. Willey and Sons, New York, 1964.

[15] I. J. Nagarath and D. P. Kothari, "Modern Power System Analysis" Tata McGraw Hill publishing company Limited, 1989.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 219-223

DISCRETE QUASI-ARMA MODELLING OF SYSTEMS WITH POINT DELAYS

by M. de la Sen, Departamento de Electricidad y Electrónica, Facultad de Ciencias, Universidad del País Vasco, Leioa (Bizkaia), Aptdo. 644 de Bilbao, SPAIN.

Indexing terms: ARMA models, Point delays.

Abstract: The letter deals with the derivation of a discrete model for linear invariant continuous systems involving point delays and subject to sampling through a zero-order hold. The used method consists of the application of the Cayley-Hamilton Theorem to the state transition matrix. The model consists of an ARMA part involving a finite set of inputs and outputs plus nonrecursive additive terms associated with initial conditions and old inputs. Such models can be relevant in parameter estimation or implementation of discrete adaptive control.

Introduction: Systems involving internal (i.e., in the state) and external (i.e., in the output) delays have received much attention in the last years [1-3]. The characterization of their typical properties like observability, controllability and stabilizability become more involved to analyze than in the delay-free case [1-3]. In particular, several (non-equivalent) definitions of observability have been given [1] and sometimes the plant stabilization can only be achieved with unstable controllers [2]. Consider the system

$$\dot{x}(t) = A\,x(t) + \sum_{i=1}^{p} A_{0i}\,x(t-h_i) + B\,u(t) + \sum_{i=1}^{q} B_{0i}\,u(t-h'_i) \qquad (1.a)$$

$$y(t) = C\,x(t) \qquad (1.b)$$

for all $t \geq 0$ where $x(.) \in \mathbb{R}^n$, $u(.) \in \mathbb{R}^m$, with $u(t) = 0$ for $t < 0$, and $y(.) \in \mathbb{R}^m$ are, respectively, the state, input and output and $h_{(.)} > 0$ and $h'_{(.)} > 0$ are point internal and external delays with $h_i < h_{i+1}$, $h_j < h'_{j+1}$; $i = 1, 2, \ldots, p-1$; $j = 1, 2, \ldots, q-1$. $x(.)$ is initialized with an arbitrary piecewise continuous function $\varphi : [-h_p, 0] \rightarrow \mathbb{R}^n$ with $\varphi(0) = x_0$. The unique solution of eqn. 1.a on $[0, \infty)$ is [3-4]

$$x(t) = e^{At}\,x_0 + \int_0^t e^{A(t-\tau)} \Big[\sum_{i=1}^{p} A_{0i}\,x(\tau - h_i) + \tilde{u}(\tau) \Big]\, d\tau \,, \quad t \geq 0 \qquad (2)$$

with $\tilde{u}(t) = B\,u(t) + \sum_{i=1}^{q} B_{oi}\,u(t-h_i')$.

Underlined State transition matrix: A unique state transition matrix $\Psi(t)$ exists for the unforced eqn. 1.a satisfying the differential system

$$\dot{\Psi}(t) = A\,\Psi(t) + \sum_{i=1}^{p} A_{oi}\,\Psi(t-h_i),\ t > 0\ ;\ \Psi(0) = I\ ,\ \Psi(t) = 0 \ \text{ for } t < 0 \tag{3}$$

which is given for $t \geq 0$ by the nonsingular matrix function

$$\Psi(t) = e^{At}\left[I + \sum_{i=1}^{p} \left(\int_{h_i}^{t} e^{-A\tau} A_{oi}\,\Psi(\tau-h_i)\,U(t-h_i)\,d\tau \right]\right. \tag{4}$$

where $U(\tau)$ is the unity step function at $\tau = 0$. Eqn. 2 can be equivalently written by using eqn. 4 as

$$x(t) = \Psi(t)\,x_0 + \sum_{i=1}^{p} \left(\int_{0}^{h_i} \Psi(t-\tau)\,A_{oi}\,\Psi(\tau-h_i)\,d\tau \right)$$

$$+ \int_0^t \Psi(t-\tau)\,\tilde{u}(\tau)\,d\tau\ ,\ t \geq 0 \tag{5}$$

Outline of proofs: Taking time-derivatives in eqn. 4 with the use of eqn. 3 proves that eqn. 4 is the unique solution of eqn. 3 by using standard unicity arguments^4. Now, take time-derivatives in eqn. 5 with the use of eqns. 3-5 to yield eqn. 1.a . This proves that eqn. 5 is its unique solution, thus being equivalent to eqn. 2 , by using the above unicity arguments.

Statement of the problem: Assume that $u(t)$ is discretised with a sampler and zero-order-hold so that $u[kT+\tau] = u_k$, all $\tau \in [0, T)$; $k = 0, 1, \ldots$ with T being the sampling period, and

$$\tilde{u}_k \triangleq \tilde{u}(kT) = \tilde{u}[kT+\tau] = B\,u_k + \sum_{j=1}^{q} B_{oj}\,u_{k-\delta_j} = \sum_{j=o}^{q_1} \bar{B}_j\,u_{k-j} \tag{6.a}$$

$$\bar{B}_o = B + \sum_{j=1}^{\bar{q}_o} B_{oj}\ ;\ \bar{B}_i = \sum_{j=\bar{q}_{i-1}}^{\bar{q}_i} B_{oj} \tag{6.b}$$

all $\tau \in [0, T)$ where δ_j (being nonnecessarily distinct) and $\bar{q}_i$ are unique nonnegative integers fulfilling $(k-\delta_j)T \leq kT-h_j' < (k-\delta_j+1)T$; $(k-i)T \leq kT-h_\ell' < (k-i+1)T$; $j = 1, 2, \ldots, q$; $\ell = \bar{q}_{i-1}, \bar{q}_{i-1}+1, \ldots, \bar{q}_i$; $i = 0, 1, \ldots, q_1$ with $q_1 = \delta_q$, $\bar{q}_{-1} = 1$ and $\bar{q}_{q_1} = q$ (if $\bar{q}_o = 0$ then $\bar{B}_o = B$). Eqns. 6 quantify the contribution of the external delays to the solution (eqns. 2 or, equivalently, 5) by locating each time $kT-h_j'$ within a sampling interval belonging to $[k_o T, kT)$ some $k_o < k$. The discrete model is derived as follows when applying the inputs eqns. 6 .

First, apply recursively eqn. 5 from $t = (k+n)T$ to $t = kT$ with initial conditions on $[kT-h_p, kT]$, together with eqns. 6 to yield:

$$x_{k+n} \triangleq x\,[(k+n)T] = \Psi^n(T)\,x_k + \sum_{i=0}^{n-1} \sum_{\ell=1}^{p} \int_0^{h_\ell} \Psi^{n-i-1}(T)\,\Psi(T-\tau)\,A_{0\ell}\,x\,[(k+i)T+\tau-h_\ell]\,d\tau + \sum_{i=0}^{n-1} \sum_{j=0}^{q_1} \left(\int_0^T \Psi^{n-i-1}(T)\,\Psi(T-\tau)\bar{B}_j\,d\tau\right) u_{k+i-j} \tag{7}$$

Apply again eqn. (5) with initial conditions on $[-h_p, 0]$ to yield:

$$x\,[(k+i)T+\tau-h_\ell] = \Psi\,[(k+i)T+\tau-h_\ell]\,x_0 + \sum_{\ell'=1}^{p} \int_0^{h_{\ell'}} \Psi\,[(k+i)T+\tau-\tau'-h_\ell]\,A_{0\ell'}\,\Psi(\tau'-h_{\ell'})\,d\tau' + \sum_{\ell'=0}^{k+i-2} \sum_{j=0}^{q_1} \left(\int_0^T \Psi^{k+i-2-\ell'}(T)\,\Psi(T-\tau')\,\bar{B}_j\,d\tau'\right) u_{\ell'-j}$$
$$+ \left(\int_0^{T+\tau-h_\ell} \Psi(T+\tau-h_\ell-\tau')\,\bar{B}_0\,d\tau'\right) u_{k+i-1} + \sum_{j=1}^{q_1} \left(\int_0^T \Psi(T-\tau')\bar{B}_j\,d\tau'\right) u_{k+i-1-j} \tag{8}$$

Eqn. 7 can be rewritten by using eqn. 8 as

$$x_{k+n} = \Psi^n(T)\,x_k + f_{k+n}(T) + g_{k+n}(T) = \Psi^n(T)\,x_k + \sum_{\ell=k-q_1}^{k+n-1} M_\ell\,u_\ell + \rho_{k+n} \tag{9}$$

where $\rho_{(.)}$, $f_{(.)}$ and $g_{(.)}$ are n-vectors defined by

$$\rho_{k+n} \triangleq f_{k+n} + \sum_{\ell=0}^{k-q_1-1} N_\ell\,u_\ell \tag{10.a}$$

$$f_{k+n} \triangleq \sum_{i=0}^{n-1} \sum_{\ell=1}^{p} \int_0^{h_i} \Psi^{n-i-1}(T)\,\Psi(T-\tau)\,A_{0\ell} \Big\{ \Psi[(k+i)T+\tau-h_\ell]\,x_0 + \sum_{\ell'=1}^{p} \left(\int_0^{h_{\ell'}} \Psi\,[(k+i)T+\tau-\tau'-h_\ell]\,A_{0\ell'}\,\Psi(\tau'-h_{\ell'})\,d\tau'\right) \Big\}\,d\tau \tag{10.b}$$

$$g_{k+n} \triangleq \sum_{\ell=k-q_1}^{k+n-1} M_\ell\,u_\ell + \sum_{\ell=0}^{k-q_1-1} N_\ell\,u_\ell = \sum_{i=0}^{n-1} \sum_{j=0}^{q_1} \left(\int_0^T \Psi^{n-i-1}(T)\Psi(T-\tau)\bar{B}_j\,d\tau\right) u_{k+i-j}$$
$$+ \sum_{\ell=1}^{p} \sum_{i=0}^{n-1} \Big\{ \sum_{\ell'=0}^{k+i-2} \sum_{j=0}^{q_1} \left[\int_0^{h_\ell}\int_0^T \Psi^{k+i-2-\ell'}(T)\,\Psi(T-\tau)\,A_{0\ell}\,\Psi(T-\tau')\,\bar{B}_j\,d\tau'\,d\tau\right] u_{\ell'-j}$$
$$+ \left(\int_0^{h_\ell}\int_0^{T+\tau-h_\ell} \Psi(T+\tau-h_\ell-\tau')\,\bar{B}_0\,d\tau'\,d\tau\right) u_{k+i-1}$$
$$+ \sum_{j=1}^{q} \left(\int_0^{h_\ell}\int_0^T \Psi(T-\tau')\,\bar{B}_j\,d\tau'\,d\tau\right) u_{k+i-1-j} \Big\} \tag{10.c}$$

where $M_{(.)}$ and $N_{(.)}$ are $n \times m$-matrix sequences which are uniquely determined from the last right-hand-side of eqn. (10.c) by separating the contributions of the input on the intervals $[0, (k-q_1-1)T]$ and $[(k-q_1)T, (k+n-1)T]$, the second one being related to the current modelling interval $[kT, (k+n)T]$ after rewritting eqn. 10.c with only a summation index for the input. For instance, the first additive term of the second right-hand-side of eqn. 10.c is identical to $\sum_{i=0}^{n-1} \sum_{j=\max(0,i-q_i)}^{i} \left(\int_0^T \Psi^{n-i-1}(T)\Psi(T-1)\,\bar{B}_{i-j}\,d\tau \right) u_{k+j}$ while similar expressions are available for the remaining additive terms of $g(.)$. Now, the application of the Cayley-Hamilton Theorem[4] to $\Psi^n(T)$ to yield $\Psi^n(T) = \sum_{j=1}^{n} \alpha_{n-j}\Psi^{n-j}(T)$ and the substitution of $\Psi^{n-j}(T)\,x_k = x_{k+n-j} - \sum_{\ell=k-q_1}^{k+n-j-1} M_\ell u_\ell - \rho_{k+n-j}$ (see eqn. 9 with the change $n \rightarrow n-j$) into eqn. 9 leads to

$$x_{k+n} = \sum_{j=1}^{n} \alpha_{n-j} x_{k+n-j} + \left[\sum_{\ell=k-q_1}^{k+n-1} M_\ell - \sum_{j=1}^{n} \sum_{\ell=k-q_1}^{k+n-j-1} \alpha_{n-j} M_\ell \right] u_\ell + \rho_{k+n} - \sum_{j=1}^{n} \alpha_{n-j}\,\rho_{k+n-j} \tag{11}$$

Finally, the substitution of eqn. 1.b into eqn. 11 yields the next quasi-ARMA (QUARMA) discrete model

$$\mathbb{A}(q^{-1})\,y_{k+n} = \mathbb{B}(q^{-1})\,u_{k+n} + \sigma_{k+n}\,, \quad k \geq 0 \tag{12}$$

where $\sigma_{(.)}$ is a signal being dependent on initial conditions and previous inputs to those within the current modelling interval $[kT, (k+n)T]$, and $\mathbb{A}(q^{-1})$ and $\mathbb{B}(q^{-1})$ are polynomials in the time-delay operator q^{-1} (i.e., $q^{-1} z_k = z_{k-1}$) whose definition follows directly from (12).

Interpretation of eqn. 12 : The reason for the presence of the signal $\sigma_{(.)}$ in eqn. 12 which makes the discrete model not to be ARMA is that the semigroup and state transition properties, which are well-known in the delay-free case[4], are not fulfilled by system eqns. 1. Thus, $\Psi(t_2) \neq \Psi(t_2-t_1)\,\Psi(t_1)$ except when $0 \leq t_1 \leq t_2 \leq h_1 \Rightarrow \Psi(t) = e^{At}$, $t \in [0, h_1]$ (see eqn. 4) and $x(t_2) \neq \Psi(t_2-t_1)\,x(t_1) + \int_{t_1}^{t_2} \Psi(t-\tau)\,\tilde{u}(\tau)\,d\tau$ (see eqn. 5). A first consequence is that a possible substitution $x_0 = \Psi^{-k}(T)\,[x_k - f_k(T) - g_k(T)]$ into eqn. 10.b to incorporate a new term from the non-ARMA contribution term σ_{k+n} of model eqn. 12 to its ARMA part would not be useful for modelling purposes since, because the semigroup property fails to hold.

References

[1] LEE, E.B. and OLBROT, A.W.: "Observability and related structural results for linear hereditary systems", Int. J. of Control, 1981, 34, pp. 1061-1078.

[2] OLBROT, A.W.: "Stabilizability, detectability and spectrum assignment for linear autonomous systems with general time delays", IEEE. Trans. Automat. Control, 1978, AC-23, pp. 887-890.

[3] DE LA SEN, M.: "Fundamental properties of linear control systems with after-effect. Part I: The continuous case", Mathematical and Computer Modelling, 1988, 10, pp. 473-489.

[4] ZADEH, L.A. and DESOER, C.A.: "Linear systems Theory" (McGraw-Hill, New York, 1963).

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 225-236

PUBLIC SECTOR AND ECONOMIC GROWTH. ESTIMATION AND SIMULATION.

Angelo Fusari
Institute of Studies for Economic Programming (ISPE),
Corso Vittorio Emanuele, 284, 00186 Rome, Italy.

Abstract.The model that this paper sets out is aimed at evaluating, at macro level, the influence of public services and public capital on economic growth. It has been estimated in continuous time, using a FIML estimator. Some forecasting and simulation experiments are carried out.

1. INTRODUCTION

There exists tight interrelationships between the efficiency of the public sector, economic growth and the public sector deficit. It is of vital importance for economic policy to highlight these interrelationships, particularly with reference to public sector activities characterized by market failure, as their inefficiencies cannot be overcome through privatization.

The distinction between public sector services received by families and firms is a basic feature of this work. It shows the interaction between these services (as well as public sector capital) and the private sector value added, capital accumulation, prices, wages and, as a consequence, public sector revenue, expenditure and debt.

In particular, the model proves that an increase in public services and capital stimulates positively the rate of growth of the private sector and helps curb inflation, public sector expenditure and deficit, thus setting in motion a virtuous circle of stabilization of the economy. Conversely,inefficiency of the public sector results in a vicious circle.

A limitation of the model structure is due to the fact that, while this specifies the effects of income distribution on prices, through the wage equation, it does not consider the effects that income distribution may have, via the rate of profit, on accumulation and the rate of growth of income. To take into account these effects, a different specification of the model would be requested. Precisely, quantities should not be explained by a production function; this should only influence them indirectly, through costs formation, while some other additional effects on production and accumulation should be displayed by demand conditions and income distribution.

A model of this kind has been exposed elsewhere[1]; it specifies the effects that the public administration inefficiencies display on the rate of growth through their impact on the private sector profit rate.

Section 2 describes the model structure and sketch out some aspect of the operative mechanism of the model. Section 3 presents the estimation results and section 4 carries out some simulation experiments aimed at evaluating the economic effect of policies centered on the improvement of public services and capital.

2. THE MODEL STRUCTURE

The Italian economy is characterized by profound disequilibria and dualism, both at sectoral and territorial level, that largely affect the relations between the public and private sectors. As a matter of fact, the public sector has largely operated as a refuge for the excess of labour; such a behaviour has slowed down its efficiency, especially in the less developed areas.

The model outlines the effects of the above tendency and wage imitation on the cost sustained by the public administration to produce services (cost mainly due to wages), then on public expenditure and deficit. Another important feature of the model is the specification of the impact displayed by the public services received by firms and public infrastructures on the productivity of the private sector. Besides, the inclusion, in the wage equation, of the trade-off between the public services received by families and wages allows us to capture some inflationary effect of the inefficiency of the public sector.

The result is that public sector efficiency and production influence growth, inflation and public sector deficit. These effects can be described as follows:

The production of public services received by firms and public investment stimulate the private sector rate of growth as well as sectoral productivity, while the increase in public services received by families curbs wage dynamics and inflation. At the same time, the reduction in the dynamics of the nominal wages paid by the public administration reduces the rise of this component of public expenditure, while the deceleration of inflation will determine a parallel deceleration of some other component of public expenditure, first of all interest paid on public debt, if we associate to that deceleration a parallel reduction of nominal interest rate. We can see, therefore, that the production and efficiency of public administration (P.A.) have a remarkable stabilizing effect on economy.

---

[1] See Fusari A. (1993). Of course, in a strictly neoclassical approach the problem sketched above takes a different form, being income distribution determined by the parameters of the production function, while accumulation is determined by saving, i.e. the preference between present and future, in the hypothesis that saving equals investment.

A disequilibrium continuous time model has been specified. It is formed by 11 stochastic equations of partial adjustment and 3 identities.

Equation 1 is a Cobb-Douglas production function explaining the private sector value added, with constant return to scale of capital and labour and increasing return with reference to the totality of its explanatory variables that also include public sector services received by firms and the lagged public investment expressing the variation of the stock of public capital, mainly infrastructure.

Equation 2 explains employment in the private sector simply through the inverse of the production function. This does not cause serious specification problem as function 2 is expressed as a first order adjustment equation, not as a zero order one. For its part, equation 3 explains dependent employment in the private sector as a share of the total private employment, which increases when the rate of growth of the private value added slows down as, in the depression, the reduction of the demand of labour stimulates independent employment.

The unit nominal wage in the private sector is expressed as a function of average labour productivity, consumer prices, the taxes payed by workmen. Besides, a trade-off between public services received by families and wages is postulated, in the event that a lack of these services increases families' spending to purchase substitutes.

Equation 5 of the unit nominal wage in the P.A. presumes that this follows the wages earned in the private sector, by imitation. In the short run, even the contrary may happen. But wage leadership by the public sector does not seem plausible in general, as the wages it pays have to financed, in the end, by taxation, so that they depend on the capacity to pay of the private sector.

Equation 6 bases the explanation of consumer prices on the unit labour cost and the rate of indirect taxes.

Public expenditure for interest on public debt is given by a linear approximation of the product between the amount of such debt and interest rate on government bonds; while the variation of public debt is expressed as a function of the public sector deficit.

Accumulation is explained through the principle of adjustment of the capital stock. Exports are expressed as a function of the ratio between domestic and international prices and world demand; while imports are a function of private value added and relative price.

Equation 12 for the value added od P.A. at current prices, expressing the labour cost soustained by P.A. for producing services, is given by an identity reproducing the way of its statistical computation. Finally, public sector deficit is the identity between public revenue and expenditure.

Model equations

$$(1)\quad DlogVAC=\alpha_1\log(\frac{\hat{VAC}}{VAC})$$

$$(1.1)\quad \hat{VAC}=\gamma_1 L_T{}^{\beta_1}K^{(1-\beta_1)}PS_i{}^{\beta_2}(\frac{I_p}{P_k})^{\beta_3}_{-4}$$

$$(2)\quad DlogL_T=\alpha_2\log(\frac{\hat{L}_T}{L_T})$$

$$(2.1)\quad \hat{L}_T=\gamma_2 VAC^{\frac{1}{\beta_1}}K^{(\frac{\beta_1-1}{\beta_1})}PS_i{}^{-\frac{\beta_2}{\beta_1}}(\frac{I_p}{P_k})^{-\frac{\beta_3}{\beta_1}}$$

$$(3)\quad DlogL_d=\alpha_3\log(\frac{\hat{L}_d}{L_d})$$

$$(3.1)\quad \hat{L}_d=\gamma_3 L_T^{\beta_4}e^{-\beta_{17}DlogVAC}$$

$$(4)\quad Dlogw_p=\alpha_4\log(\frac{\hat{w}_p}{w_p})$$

$$(4.1)\quad \hat{w}_p=\gamma_4(\frac{VAC}{L_T})^{\beta_5}PC^{\beta_6}CSIR^{\beta_7}PS_1{}^{-\beta_8}$$

$$(5)\quad Dlogw_{PA}=\alpha_5\log(\frac{\hat{w}_{PA}}{w_{PA}})$$

$$(5.1)\quad \hat{w}_{PA}=\gamma_5 w_p{}^{\beta_9}$$

$$(6)\quad DlogPC=\alpha_6\log(\frac{\hat{PC}}{PC})$$

$$(6.1)\quad \hat{PC}=\gamma_6(\frac{w_p}{\frac{VAC}{L_T}})^{\beta_{10}}(\frac{T_i}{VACP})^{\beta_{11}}$$

$$(7)\quad logRC=C_1\beta_{12}logB+C_2\beta_{13}i+\log\gamma_7$$

$$(8)\quad DlogB=\alpha_7\log(\frac{\hat{B}}{B})$$

$$(8.1)\quad \hat{B}=\gamma_8 DP^{\beta_{14}}$$

$$(9)\quad Dlogk=\alpha_8\log(\frac{\hat{k}}{k})$$

$$(9.1)\quad \hat{k}=\gamma_9 VAC^{\beta_{15}}K^{-\beta_{16}}$$

$$(10)\,DlogEXP=\alpha_9\log(\frac{\hat{EXP}}{EXP})$$

$$(10.1)\,\hat{EXP}=\gamma_{10}PC^{-\beta_{18}}(P_i\$)^{\beta_{19}}Y_w^{\beta_{20}}$$

$$(11)\,DlogIMP=\alpha_{10}\log(\frac{\hat{IMP}}{IMP})$$

$$(11.1)\,\hat{IMP}=\gamma_{11}VAC^{\beta_{21}}(\frac{PC}{PIMP\$})^{\beta_{22}}$$

$$(12)\quad VAPAP=w_{PA}*L_{PA}$$

$$(13)\quad DlogK=k$$

$$(14)\quad DP=(VAPAP+RC+I_p+RU-CSIR-T_i-RE)$$

Endogenous

VAC = Private sector value added in real term.
L_T = Total employment in the private sector.
L_d = Dependent employment in the private sector.
w_p = Money wage rate in the private sector.
W_{PA} = Money wage rate in P.A.
PC = Consumer price index.
RC = Public sector interest's payments.
B = Total amount of public debt.
EXP = Real exports of goods and services.
IMP = Real imports of goods and services.
VAPAP=Nominal value added of P.A.

DP = Public sector deficit.
K = Real capital stock in the private sector.
k = Proportional rate of increase of K

Exogenous

PS_i = Public services received by firms.
PS_l = Public services received by families.
I_p = Public sector investment in current lire, after deduction of that in housing.
P_k = Price index of capital goods.
P = Deflator of the private sector value added.
P_i = Exports unit value index of manufactures of OCSE countries, in US dollars.
PIMP = Price index of imports, in US dollars.
\$ = Exchange rate lira/dollar.
Y_w = World real income.
i = Interest rate of government bonds.
L_{PA} = Employment in P.A.
T_i = Indirect taxes.
CSIR = Social security contribution plus taxes paid by workmen.
RE = Remaining public sector revenue.
RU = Remaining public sector expenditure.
α = Adjustment parameter.
β = Elasticity.
γ = Intercept.
D = Symbol of derivative.
$\hat{}$ = Symbol of partial equilibrium.

3. ESTIMATION RESULTS

The estimation is in continuous time; it utilizes a FIML estimator and quarterly data series ranging from 1980 to 1991. The estimated parameters have an asympotic normal distribution.

Table 1

Estimated adjustment parameters

Parameter	Estimate of parameter	Asymptotic standard err.	t-ratio	Mean time lag(in quar.)
α_1	0,640	0,144	4,43	1,562
α_2	0,025	0,020	1,22	40.0
α_3	0,248	0,051	4,89	4,032
α_4	1,769	0,422	4,19	0,565
α_5	0,914	0,272	3,45	1,063
α_6	0,157	0,036	4,37	6,369
α_7	0,035	0,012	2,94	28,571
α_8	0,461	0,073	6,35	2,169
α_9	3,109	0,840	3,70	0,322
α_{10}	1,756	0,445	3,95	0,569

Table 2

Estimated elasticities and intercepts

Parameter	Enter.equat.numb	Estim. of par.	Asympt. stand. err.	t-ratio
β_1	(1)	0,632	0,072	8,74
β_2	(1)	0,410	0,109	3,76
β_3	(1)	0,325	0,049	6,67
β_4	(3)	0,990	0,138	7,16
β_5	(4)	0,567	0,057	9,95
β_6	(4)	0,475	0,063	7,49
β_7	(4)	0,411	0,051	8,10
β_8	(4)	0,691	0,244	2,83
β_9	(5)	1,060	0,029	36,85
β_{10}	(6)	0,644	0,109	5,91
β_{11}	(6)	0,435	0,130	3,35
β_{12}	(7)	0,736	0,223	3,30
β_{13}	(7)	0,075	0,037	2,02
β_{14}	(8)	0,707	0,329	2,15
β_{15}	(9)	0,024	0,002	10,91
β_{16}	(9)	0,025	0,002	11,95
β_{17}	(3)	0,951	0,367	2,60
β_{18}	(10)	0,302	0,085	3,56
β_{19}	(10)	0,167	0,069	2,44
β_{20}	(10)	2,018	0,155	12,98
β_{21}	(11)	2,054	0,115	17,86
β_{22}	(11)	-0,062	0,068	0,91
$\log\gamma_1$	(1)	6,267	0,577	10,85
$\log\gamma_2$	(2)	-9,890	0,712	13,89
$\log\gamma_3$	(3)	0,362	1,364	0,27
$\log\gamma_4$	(4)	-0,097	0,008	11,77
$\log\gamma_5$	(5)	-0,813	0,150	5,40
$\log\gamma_6$	(6)	-0,120	0,055	2,19
$\log\gamma_7$	(7)	-1,247	0,261	4,78
$\log\gamma_8$	(8)	-0,0002	0,013	0,02
$\log\gamma_9$	(9)	-7,483	3,697	2,02
$\log\gamma_{10}$	(10)	7,654	1,413	5,42
$\log\gamma_{11}$	(11)	13,832	1,373	10,07

$C_1 = 1.38413$ $C_2 = 37.4898$.

The model being expressed in logarithms, the estimated parameters give the elasticities. All parameters are significantly different from zero at the 1 percent level, except one and two intercepts.

It is important to consider the value of the parameters expressing the effect of the public services and public capital on the value added and wages of the private sector. At this purpose, we can see that an increase of 1 percent in public services received by firms cause (through β_2) a 0.41 percent increase in the private sector value added; for its part, an increase of 1 percent of public sector investment in real terms cause (through β_3) a 0.33 percent increase in the private sector value added, while an increase of 1 percent of public services received by families slows inflation of 0,69 percent (through β_8). This proves the relevance of the public sector activity to economic growth and inflation and, as a consequence, to public sector expenditure, revenue and public debt.

The model structure and the estimation results seem to refuse the Ricardian equivalence between taxes and public debt. As a matter of fact, we can see that a reduction in public sector deficit by increasing taxation stimulates inflation through parameters β_7 and β_{11}. Public sector deficit seems to have lowered the impact on prices of the inefficiencies of public administration and favoured the acceptance of them by people, with respect to the case that they had implied a higher taxation. In addition, the inefficiencies of the public administration reduce the rate of growth of value added both directly through the production function and owing to inflation they stimulate, that reduce (in the presence of fixed exchange rate) exports, thus causing a squeeze of domestic demand aimed at meeting the disequilibria of foreign accounts. As is well known, a rate of growth inferior to the real interest rate implies the unsustainability of public debt. For this purpose, it is worthwhile pointing out that public sector debt caused by some persistent inefficiencies of public administration largely differs from the one due to exceptional needs (in Ricardo's sense). At first, it hides the above inefficiencies, that therefore can freely develop; but a cumulative and unrestrainable growth of public sector deficit will follow, pushed by such inefficiencies.

4. ON THE EFFICIENCY OF FORECASTS AND SOME SIMULATION EXPERIMENTS.

A picture of the forecasting efficiency of the model may be given by the root mean square errors (RMSE) resulting both from static and dynamic simulations. Being the model expressed in logarithms, the RMSE give the percent level of the average errors of the endogenous variables.

Table 3

Ex post root mean-square errors (RMSE) of forecasts (1980/91).

Variable	Single period forecast	Dynamic forecast.
VAC	0.0078	0.0108
L_T	0.0026	0.0067
L_D	0.0035	0.0084
w_P	0.0125	0.0165
w_{PA}	0.0301	0.0443
PC	0.0035	0.0092
RC	0.0176	0.0852
B	0.0125	0.0308
k	0.0001	0.0003
EXP	0.0250	0.0252
IMP	0.0211	0.0275
VAPAP	0.0245	0.0518
K	0.0	0.0019
DP	0.0325	0.0890

We can see that the static simulation mostly shows errors lower than 1 percent, while the dynamic simulation shows errors a little higher; but public debt and interest show an 8 percent error. These results can be considered quite good. We now turn to some simulation experiment. Simulation (a) hypothizes a rate of growth of the P.A.real value added received by firms, as well as of public sector investment, 1 per cent higher per year than the actual one (this may be a reasonable assumption, as the efficiency of the Italian P.A. is very low). Besides, it hypotizes that the real interest rate remains unchanged with respect to the actual one, i.e. the decrease in inflation rate causes a parallel decrease in nominal interest rate.

Simulation (b) adds one assumption to the previous one : the yearly rate of growth of public sector services received by families is a 1 percent higher than the current one.

Simulation experiments also suppose that price variations influence P.A. revenues and expenditure, exogenous in the model. Finally, it is assumed that the increase in the value added of P.A. results from labour productivity increases in this sector, achieved at zero cost. This assumption is not plausible. However, we will see that simulation experiments produce a considerable reduction in the cost sustained by P.A. for producing services. This means that some cost eventually requested for implementing the efficiency of P.A. could be met by the above reduction.

Table 4 below shows the results of the simulation experiments, expressed in liras or units of labour. They refer to the final year (1991) of the sample period, thus expressing the cumulated effects, during 11 years, of the hypotheses formulated by simulations (a) and (b). These results will be compared to the actual values taken by the model variables in 1991.

Table 4. 1991 values, in billion of lire; units of labour in thousands. w_r and w_{PAr} are real values of w_p and w_{PA}

Variable	Actual	Tendential	Simul. (a)	Simul. (b)
VAC	753491	751996	855370	885067
L_T	19787	19819	19990	20036
w_p	37.7	37.5	36.2	31.5
w_r	26.4	26.3	29.0	28.1
w_{PA}	49.1	47.5	45.9	39.8
w_{PAr}	34.6	33.3	36.8	35.5
PC	1.426	1.425	1.248	1.122
RC	145550	147635	111817	77093
B	1446360	1503069	1497353	1348562
EXP	222983	216510	224468	231135
IMP	264338	261098	289758	300048
VAPAP	184666	180591	174892	151890
K	4163937	4160725	4343322	4386841
DP	146038	144197	131331	93806
DP/VACP	0.14	0.138	0.111	0.076
B/VACP	1.387	1.4436	1.264	1.1

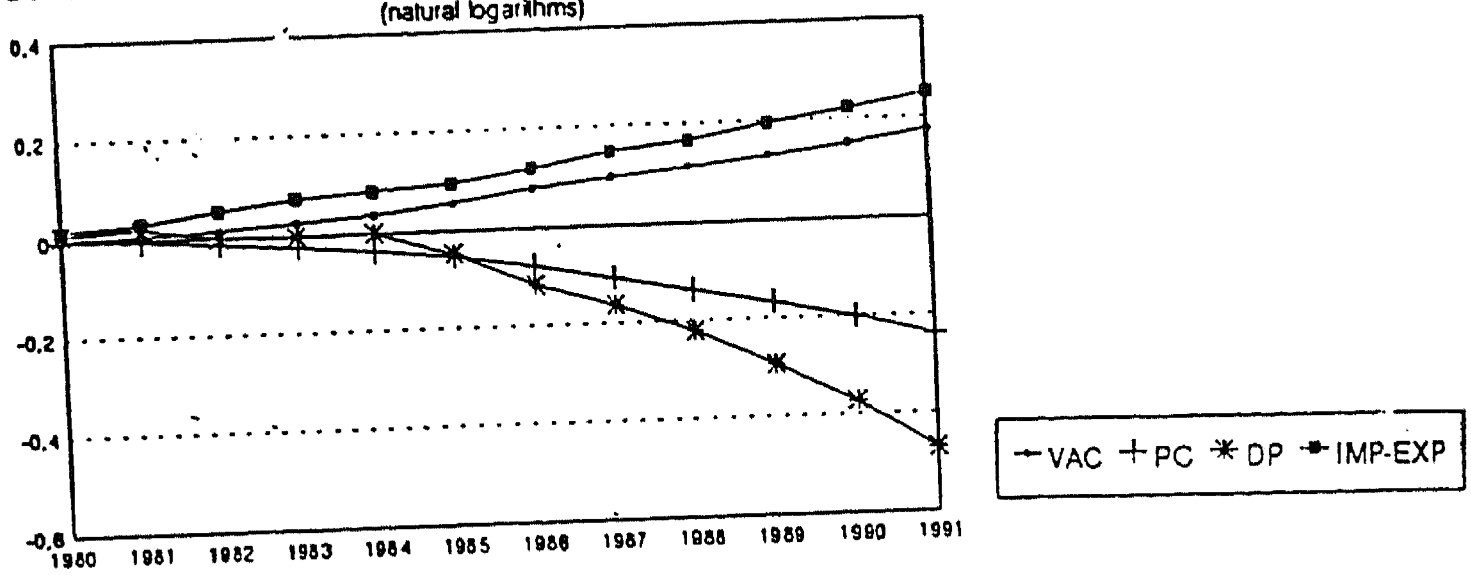
As we can see, the values obtained from the neutral solution are very close to the actual ones; the differences between the two are mostly inferior to 1 percent, while the largest of all is less than 4 percent. For a better appreciation of these results, it is important to keep in mind that they describe the final year of the eleven years trajectory.

Also impressing is the remarkable increase in the value added of the private sector resulting from simulation (a) and (b), together with the decrease in inflation rate and the strong reduction in public sector deficit and public debt. It is also worthwhile pointing out the substantial decrease, in both simulations, of the cost sustained by P.A. for producing services. The rate of growth of employment in the private sector appears to be very low. It seems evident that the achievement of a substantial increase in employment requires a reduction in the hours worked per head.

Note that the simulations results also show some increase in the real wages (with respect to the tendential solution). This effect, together with the increase in public sector services received by families that the simulations assume, implies a remarkable increase in social welfare, thus implementing the possibility to carry out income policy.

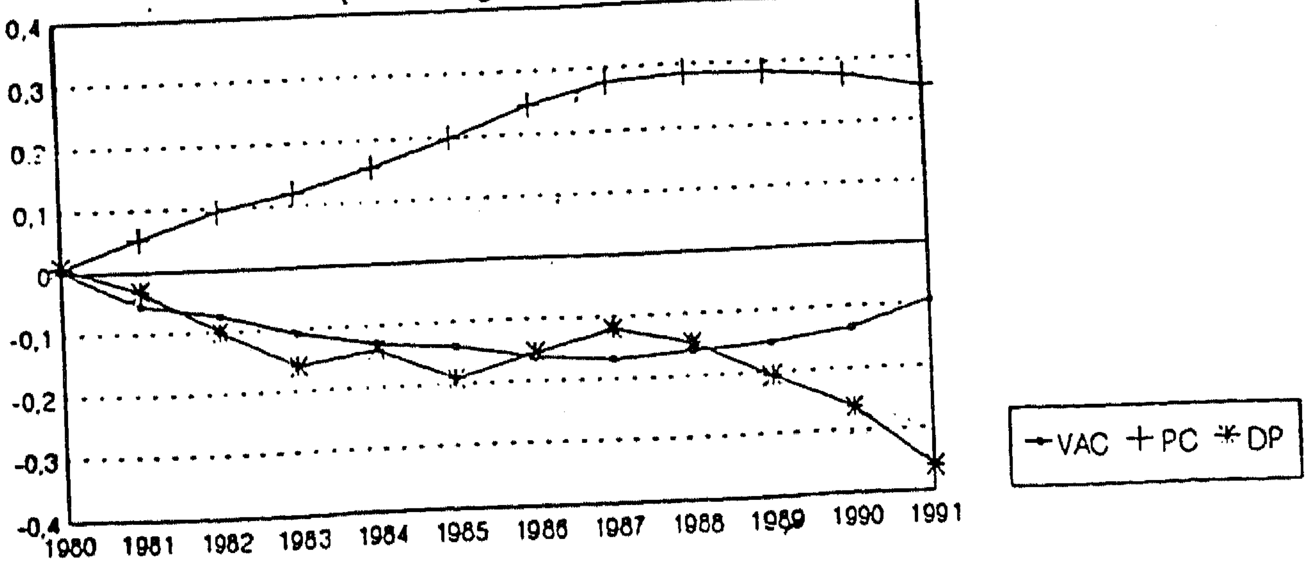
The strong reduction in public sector deficit is largely due to the decrease in interest paid on public debt. This aspect deserve some exploration with reference to its linkage with foreign sector accounts. From table 4 above we can see that simulations imply a remarkable increase in the deficit of the balance of payments, in the hypothesis that exchange rate is exogenous, i.e. this does not vary with respect to the actual one; such a deficit renders difficult a reduction in nominal interest rate (hence public debt) after the reduction in inflation rate, as the equilibrium of foreign accounts requires a high interest rate to attract capital from abroad. This evidentiates the role of the foreign account obstacle to growth, due to the high elasticity of imports with respect to private value added, and the domestic one due to the inefficiencies of public sector. A better clarification of this aspect is given by the figure below.

Deviations of the results of simulation (b) from the tendential solution.
(natural logarithms)



For a further deepening of this question, a control problem has been formulated, which takes public sector and foreign accounts deficits as targets, and public sector services and investment as instruments. More precisely, it postulates, for the final period, a public debt growth equal to that of real private sector value added plus inflation, and a convergence of foreign accounts deficit towards some reasonable value. The control variables choose the elasticity of public sector services received by firms and public sector investment with respect to the actual value of these variables, but for the public sector services received by families a proportional relation with the real private sector value added has been postulated.

Deviations of the result of the control experiment from the tendential one.
(natural logarithms)



As we can see, the simultaneous achievement of the above targets implies a substantial deceleration in the rate of growth of the private sector value added (with respect to the neutral solution), accompanied by a remarkable acceleration of inflation due to the decrease in the dynamics of labour productivity. To avoid these inconsistencies, a reduction in the elasticity of imports with respect to the private sector value added appears to be necessary.

5. CONCLUSION

This essay shows that it is mistaken to pursue the compression of the public sector deficit through the reduction of public sector services and investment, as this will cause, in the medium run, adverse effects. It is evident that the increase in the efficiency of P.A. represents a leading road to stimulate growth and reduce inflation and public debt.

A further deterioration of public sector services vital for firms obstructs the activity of these. Likewise, the reduction of public sector services received by families compresses social welfare. This sharpen the conflict for income distribution and inflation, makes difficult wage income, curbs accumulation and growth.

The widespread inefficiencies of the Italian P.A. are not unavoidable. One can find public sector services operating satisfactorily. These cases should be accurately studied, in order to understand the factors on which their success depend. But this is another story.

REFERENCES

1. Aschauer, D.A. (1989), "Is public expenditure productive?", Journal of Monetary Economics.

2. Barro, R.J.(1990), "Government spending in a simple model of endogenous growth", Journal of Political Economy, October.

3. Fusari, A. (1990), "A macrodynamic model centered on the interrelationships between the public and private sectors of the Italian economy", New York Economic Review,Spring.

4. Fusari, A. (1993) "A cyclical model of economic growth: Theory and application", presented at the Far Eastern Meeting of the Econometric Society, Taipei, June 24-28.

5. Kormendi, R.C. (1983) "Government debt, government spending and private sector behavior", The American Economic Review.

6. Romer, D. (1988), "What are the costs of excessive deficits?", NBER Macroeconomic Annual.

7. Wymer, C.R., Resimul and Apredic manuals and programs for estimation and simulation.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 237-251

AN INFORMATION SYSTEM FOR LIBRARY MANAGEMENT.
A SYSTEMS STUDY AND DESIGN

K. J. Satao
Assistant Professor
Deptt. of Computer Sc. & Engg.,
B. N. College of Engineering,
PUSAD - 445 215(INDIA)

Abstract

The information needs of human beings can include just about any subject viz. persons, places, things, methods, legal requirements, and so forth. It also can encompass many different types of materials viz. periodicals, articles, books, news papers, patents and so forth. While much of this information may be internal to the firm or the area under study, more may be located via the library.Today most of the larger corporations recognize that the library is an expensive but a vital resource. It becomes the wasted resource, if it is not used by the persons for whom it is maintained. A library may contain periodicals, magazines, books, news papers, manufacturer's literature, pamphlets, brochures, maps & atlases, directories, annual reports, business organization publications, Govt. documents, technical reports etc. Today computers can be used in just about every library activity viz. acquisition, circulation, cataloging, literature searching and so on. How can this be done is discussed in this paper.

1.0 Introduction

The concept of "management" as such deals with the activities viz. planning, organizing, directing, staffing, and controlling the resources such as man power, money, materials, machines, and facilities. To facilitate the management to take proper decisions at proper times and situations, we need what is called as an "information system". The word "system" can be described as a set of elements joined together for a common objective, whereas "information" consists of data that have been retrieved , processed, or otherwise used for the informative or interference purposes, arguments, or as a basis for forecasting or decision making.

An information system consists of the following basic activities viz.

1> Entering data into the system,
2> Processing the data (rearranging input data, computations, comparisons, and processing files),
3> Maintaining files and records,
4> Developing procedures that tell what are the data needed ? and when ?, where are they obtained from ?, and how are they used ?, as well as providing instruction routines for the processor to follow , and
5> Preparing report output.

2.0 A System's Study : The Manual System

I consider an example of a library of a typical educational institute having about 10,000 books and 1500 members consisting of students, meritorious students, teachers, H.O.D.s, non teaching staff members and the Principal. It is assumed that each student gets two library tickets, each meritorious student gets 8 library tickets(through "book bank" scheme),each teacher gets 10 library tickets, each H.O.D. gets 12 library tickets, each non teaching staff member gets two library tickets, and the Principal gets 15 library tickets. The students can get only the books issued, whereas the staff can get books/periodicals etc. The library charges 50 paise(currency) per day for books not returned within 15 days by the student members. The staff members can retain the books for the entire semester/year since they require it for the reference purposes, whereas the non-books viz. magazines have to be returned by them within 15 days. But normally no fine is levied on them. The fine is calculated and the students are asked to pay the fine in the account's department and produce the receipt. Files are maintained for the fines' receipts, special permissions from the Principal, publisher letters,member complaints & suggestions etc.and registers are maintained for the issues/receives, department wise books etc.. The transactions are written date wise in these registers for the student members. For the teaching staff members a separate register is maintained in which the transactions are entered member wise, separately, on the different pages maintained against their names.Whenever a book is issued to a member, one of his library tickets is taken by the library. At the clearance time a member has to produce all his library tickets. The library purchases minimum two books of a type. One copy, out of these two, is put in the reference section. This copy cannot be issued,normally, to a member. But with a special permission, under exceptional circumstances, from the Principal this copy can also be issued. The library also has a "Book Bank" scheme. Under this scheme, the first 10 students of a class are provided one full set of the books for the entire session. These books have to be returned at the end of the session. These books cannot, also, be issued to a normal member. But the Principal can give a special permission. The Principal can also give a special permission to waive off a fine to a particular student.

The library has a facility to reserve the books, give suggestions for purchasing the new books. The catalogs are asked and the orders are placed. The library maintains the data about the non-book literature viz. periodicals, pamphlets, brochures, maps, atlases, directories, annual reports, technical reports etc. The news papers and recreation magazines are preserved for a period of about six months and then disposed off(sell out as rough). All the technical manuals, periodicals are preserved separately for one year, and after that bound in a single volume. The library maintains author wise and title wise index of the books. The library does not have a facility of overdues monitoring except at the end of the semester or the book is very much in demand. The library does not send letters to members.

Whenever a new book is received by the library, it is entered in the general register and in the classified(department wise) register. If the book is lost, it's cost is recovered from the member responsible for losing it along with the penalty. A physical inventory is made once in a year and the cost of the missing books is recovered from the employees concerned.

The index card of each book contains the fields viz. Accession number, classification code(showing the area of the book and the shelf in which the book is stored), title of the book, names of the authors, price of the book, publisher of the book.

Each student member is given an identity card containing...

- His/her photograph,
- Membership no.,
- Full name,
- Department,
- Address,
- Specimen signature,
- Principal's signature.

The photograph and the signature are for the manual check.
Each student member has to pay a refundable library deposit of say rupees 200(currency) and the library fee of rupees 50 per year. The library ticket i.e. the B.T.(Borrower's Ticket) contains the fields viz.

-B.T. no., Year, Name, Department, Borrower's sign, Librarian's sign.

Each book in the library contains a card containing the following information...

-Classification no., Accession no., Title, Author, Borrower's name, Borrower's signature, Date of issue, Date of return, Receiver's initials.

The library ticket(B.T.) contains a plastic jacket into which the book's card will be inserted and the library ticket will be preserved by the library till the book is received back by the library. There is a separate issue counter and a separate receive counter in the library. The library has a scheme of Gate Pass for all the books issued by the library and the personal books cannot be taken inside the library.

3.0 The System's Analysis : A Proposed new System

3.1 Need for Computerization

This is because of some of the inadequacies of the present manual system viz.

1> The existing system is slow in responding to requests for information e.g., it takes long time to answer the queries viz.
  a> To whom a particular book is issued ?
  b> Whether so and so book is available in the library or not, if yes can it be issued ?
  c> What are the books issued to a particular member ?

2> The existing system cannot handle the work load e.g. there is usually too much of crowding at the issue / receive counter. Due to this heavy work load, it is frequently necessary to receive books without checking fines.

3> The existing system has problems of accuracy e.g. updated catalogs are not available, so students have to physically go into the library, find out the books of their interest and get them issued. In this process some books get stolen, and the staff has to pay hundreds of rupees for this loss.

3.2 Expectations from the Computer

A> Queries about Books/Journals/Periodicals etc.

- Is a specific book/journal/periodical available in the library ?
- If yes, can it be issued ? or to whom is it issued ? or
- Is it a reference book or a "book bank" book or is the book reserved ?
- How many copies of a particular book are available in the library ?
- How many books of a particular author are available in the library ?
- List of books/journals/periodicals in some area / department with cost.
- List of all books/journals/periodicals for physical inventory.
- Given an accession no.,give the details of the book(issued/ available / reference book/"book bank" book,if issued then to whom is it issued ?).
- Whether a particular book is overdue or how many books are overdue as of a particular date ? etc.

B> Queries about members

- Given a particular member number how many and what are the books / non-books issued to him/her, and when ?
- List of all members, department wise, who have returned all the books.
- List of all members, department wise, who have books/non-books issued.
- List of all members, department wise, who have overdue books etc.

C> Operations

- Add a new book/non-book to the library.
- Add a new member to the library.
- Delete or update a book/non-book.
- Delete or update a member.
- Issue/receive a book/non-book.
- If a book is overdue from a member then send a letter to him/her.
- Allow the member to reserve a book and give suggestions etc..

3.3 Proposal of a Computer

At this stage we have become familiar with the library data, it's requirements, and the existing manual procedures. Our next task is to do the cost/benefit analysis for using the computer. If the benefits can be quantified, the comparisons would become easy. Unfortunately, it is difficult to measure (in terms of money) the improvements in the library services due to the introduction of the computer. The benefits may be mostly intangible i.e. in terms of the members' satisfaction, less trouble to the library staff, and better maintenance. Sometimes the computerization may become a prestige point of view and the library may be computerized.

For the above system of 10,000 books,1500 members and the above expectations the following Hardware and Software will be required...

- Computer machine PC AT 386 type.
- RAM 4 MB.
- Winchester disk (80 MB).
- Cartridge tape drive(60 MB).
- Floppy disk drive(3.5").
- O. S. (multiuser, multitasking).
- Console + one terminal connected with the machine.
- Printer 132 Col, 300 CPS, DMP.
- Data base systems s/w viz. dBase III plus/ Fox Base/ UNIFY or a compiler viz COBOL and a word processing package viz. Word Star.
- Stabilizer, air conditioners, U. P. S. etc.
- 10X10 room with glass walls,two windows, two tables and two chairs etc.

This proposal will take care of the future demand of about five years. The day to day basis operating tasks can be handled by two operators and for the programming a computer consultant can be contacted who will develop the system and charge a one time cost or a computer programmer cum systems analyst can be employed. In general a single terminal may satisfy the need, but at times two terminals may be required. The introduction of the two operators will take care of their personal leaves of all the types. It is assumed that on any fine morning at least one of the two operators will be available for service. The existing library staff can also be trained for the operations and programming instead of employing new staff.

3.4 Model of the Computerized System

3.41 Initial Preparations

1> The scheme of library tickets will not be dislodged in the computerized system because in case there is any problem with the computer system viz. hardware failure or the power failure, we should be able to continue the services of the library with the library tickets.Each student of the college will be given exactly 2 B. T.s, each teacher will be given exactly 10 B. T.s, each meritorious student will be given exactly 8 B.T.s,each non teaching staff member will be given exactly 5 B.T.s, each H.O.D. will be given exactly 12 B.T.s, and the Principal will be given exactly 15 B. T.s.

2> For the convenience of the members a latest list of the books available in the library can be made available in front of the window. The entire list of the books can be printed, department wise, every month and put ready for the members reference/we may use a box file for each department and in this box file each new list of the books added to the library can be inserted department wise. It is suggested that every new book received by the library should be put in the separate display section, maintained in the library, for about a week so that the members will know the new entries of the books in the library.

3> A bunch of issue slips containing the following information will be made available(cyclostyled) to the members in front of the counter.

a> Member number :
b> Member name :
c> Accession number of the book :
d> Title of the book :
e> Signature of the member :

On the reverse of this issue slip we will have the entries...

Received the book,

Signature of the member.

A bunch of suggestion forms, containing the information as that of the Purchase file, and the reservation slips, containing the information as that of the issue slips, will also be made available(cyclostyled) to the members at the counter.

3.42 Main Front Office Jobs

1> A member who wants to have a transaction with the library, will stand in a queue in front of the window where a computer terminal is placed. If a member wants to get a book/non-book issued, he will fill a issue slip and give it at the counter along with his B.T..If a member wants to give the book back to the library, he will just give the book at the counter along with a fine receipt or a special permission, if any.

2> For the issue slip, the operator will check the signature on the B.T. and the issue slip. If they match, he will use the Transaction Entry program and check whether the book can be issued, otherwise the transaction will be rejected. In case there is a reservation for the book, the transaction will be rejected. If the book is a reference book or the "book bank" book and is available in the library but the member does not have a special permission then the operator will turn down the request and ask the member to bring a special permission. Otherwise the operator will process the transaction. If the book is not the reference book or the "book bank" book but the particular book is issued then the operator will enquire on the title of the book and will see whether another book of the same title is available, if yes then the operator will ask the member to modify the issue slip and will take the issue slip from him. In case of any of the books on that title is not available then the operator will tell the member that no such copy is available and will ask the member to reserve the book and this transaction will be over. In case the original book itself can be issued then the operator will take the issue slip from the member, write the shelf no.of the book from the classification no. of the book's entry in the computer's file on the issue slip,give a token to the member after writing the token no. on the slip,and ask the member to stand in another queue or come later.The operator will then hand over the issue slip to the library attendant to search and get the book from the shelf. Till the time the book is traced , the operator will consider the requests of the other members.After the book is received,it will be first given to the operator, the operator will make an entry into the computer. This entry will be done by the program into the Books file, the Members file, the Transactions file, and the Control file. Then the book will be given to the member after getting his signature on the reverse of the issue slip and after taking one of his B.T.'s and this transaction will be over.The issue slip and the B.T. will be preserved by the library,till the book is received back by the library. There is no provision for the Gate Pass.

3> If the member wants to give a book back to the library, the operator will check whether the book is being returned within time or not (within 15 days), from the fine-amount of the book from the computer's file. If the book is not being received within time, for the student member, the operator will ask the student to bring the receipt of the appropriate amount from the account's section and will accept the book only after the receipt or a special permission from the Principal. If the student brings the book well in time(fine-amount = 0) or along with the fine receipt or a special permission(in case of fine-amount not equal to zero) then the book will be taken by the operator and the program will make the appropriate entries in the Members file, the Books file, the Transactions file, and the Control file. The entries will be made by the Transaction Entry program and the transaction will be over.

3.43 Main Back Office Jobs

a> Day Book Printing : This will list all the transactions of the date.
b> Overdues Printing : This will find all the overdue books,write letters, and execute the Fine Calculations program.
c> Control Report Printing : This will print counters for all the 13 types of transactions. The reports in a> and c> have to be manually checked by the operators.
d> Taking back-up : Take the back-up of the Members file, Books file, the Transactions file, and the Control file etc.

4.0 The System's Design : File Design & Program Specifications

4.1 File Design

The file design phase is a very important phase in the design of a computerized application. It controls not only the type of the processing(on line or batch or real-time), but also the overall efficiency and ease of writing programs. I discuss here the file design in the COBOL environment. For our application we need the following seven files...

4.11 Members file

In our application we have 6 types of members, students (getting 2 books), teachers(getting 10 books), H. O. D.s(getting 12 books), non teaching staff members(getting 5 books), meritorious student(getting 8 books), and the Principal(getting 15 books). Let us codify these as S - student, M - Meritorious student, T - teacher, H - H. O. D., N - non teaching staff member, and P - Principal. Now our file will be a variable record length type of file having six different types of records in the same file. These six different records of the file can be defined at six different 01 levels in the FD clause. A typical 01 entry - let us say for the student will be consisting of the following fields...

```
01> MEMBER-NO                PIC 9(4). (PRIMARY KEY)
02> MEMBER-TYPE              PIC X.    (S/M/T/H/N/P)
03> MEMBER-NAME              PIC X(30).
04> MEMBER-DEPT              PIC X.
    (P-PHYSICS, C-CHEMISTRY, M-MATHS, U-ELECTRONICS, R-COMPUTER, I-CIVIL,
    M-MECHANICAL, S-SPORTS, L-LIBRARY, T-TRAINING & PLACEMENT,O-OFFICE etc.)
05> MEMBER-ADDRESS           PIC X(80).
06> NO-OF-BOOKS-GOT-ISSUED   PIC 9(2).
07> FINE-ALREADY-PAID        PIC 9(4)V99.
08> FINE-TO-BE-PAID          PIC 9(4)V99.
09> ISSUED-BOOK-ACNO-1       PIC 9(5).          |These fields will be 16 for M,
10> DATE-1                   PIC X(6). (YYDDMM) |20 for T,24 for H, 10 for N, &
11> ISSUED-BOOK-ACNO-2       PIC 9(5).          |30 for P.Thus the file will be
12> DATE-2                   PIC X(6). (YYMMDD) |of a variable record length.
```

The organization of the file will be indexed and the accessing mode will be dynamic. The fields 2, 3, and 4 will be the secondary keys.

4.12 Books file

In the library we have two types of books viz. 1. Books and 2. Non-books e.g. periodicals, journals, magazines, maps, atlases, directories, technical reports, pamphlets, brochures etc. The teachers can get both the books and the non-books issued, whereas the students can get only the books issued. Let us codify these items as...

B - NORMAL BOOK, N - NON-BOOK, R - REFERENCE BOOK , O - "BOOK BANK" BOOK.

Some of the books can be the reference books and the "book bank" books. But the majority of the books will be the normal books which could be issued to any member,if available. This file will have two 01 levels(one for B, R, and O type of the book, and the other for the N type of the book). The file organization will be indexed and the accessing mode will be dynamic. The book 01 level will have the following fields...

```
01> ACCESSION-NO          PIC 9(5). (PRIMARY KEY)
02> BOOK-TYPE             PIC X.(B/N/R/O)
03> CLASSIFICATION        PIC X(10). (CONTAINING AREA OF THE BOOK & SHELF NO.)
04> TITLE-OF-THE-BOOK     PIC X(80). (IF REQUIRED CURTAIL TO 80 CHARACTERS
05> AUTHOR-NAME-1         PIC X(30).
06> AUTHOR-NAME-2         PIC X(30).
07> AUTHOR-NAME-3         PIC X(30).
08> PUBLISHER-NAME        PIC X(30).
09> PUBLISHER-ADDRESS     PIC X(80).
10> COST-OF-THE-BOOK      PIC 9(5)V99. (IN A SINGLE TYPE OF CURRENCY ONLY)
11> VOLUME-NO             PIC 99.
12> EDITION               PIC 99.
13> YEAR-OF-PUBLICATION   PIC X(4).
14> DATE-OF-PURCHASE      PIC X(6). (YYMMDD)
15> DEPT-CODE             PIC X. (AS IN THE MEMBERS FILE)
16> BOOK-LANGUAGE         PIC X(10).
17> STATUS-OF-THE-BOOK    PIC 9.
    (0-BOOK IS AVAILABLE,1-BOOK IS ISSUED,2-BOOK IS LOST)
18> ISSUED-TO-MEMBER-NO   PIC 9(4).
19> ISSUED-ON-DATE        PIC X(6). (YYMMDD)
20> DATE-OF-RECEIVING     PIC X(6). (YYMMDD)
21> FINE-AMOUNT           PIC 9(5).
22> FINE-PAID             PIC X. (Y/N)
23> OVERDUE-LETTER-SENT   PIC X.
24> SPECIAL-PERMISSION    PIC X. (Y/N)
25> RESERVED-BY-MEM-1     PIC 9(4).
26> DATE-1                PIC X(6). (YYMMDD)
27> RESERVED-BY-MEM-2     PIC 9(4).
28> DATE-2                PIC X(6). (YYMMDD)
29> REMARKS               PIC X(15).
```

The fields 2, 4, 5, 6, 7, 8, 15, 16, 17 will all be the secondary keys. The non-book 01 level will not have the above 5,6,7,12, and 13 fields, whereas the following two fields will be added... DATE-OF-PUBLICATION PIC X(6) (YYMMDD), MONTH-OF-PUBLICATION PIC XXX(JAN/FEB/.../DEC).

4.13 Purchase file

The library members can suggest some useful books/non-books for the library. Their suggestions will be received as and when they come and will be stored into this computer file. The suggestions will be accumulated for about 15 days or one month and after this period a printout of this file will be taken and the file will be purged(file will exist but the inside data will be removed) or deleted, and manually a decision will be taken as to which suggestions are useful. The Purchase file can have the following fields...

01> TYPE-OF-THE-BOOK	PIC X. (B - BOOK, N - NON-BOOK)
02> TITLE-OF-THE-BOOK	PIC X(80).
03> AUTHOR-NAMES	PIC X(90).
04> PUBLISHER-NAME	PIC X(30).
05> PUBLISHER-ADDRESS	PIC X(75).
06> COST	PIC 9(5)V99.
07> VOLUME-NO	PIC 99.
08> EDITION	PIC 99.
09> YEAR-OF-PUBLICATION	PIC X(4). (YYYY)
10> DATE-OF-PUBLICATION	PIC X(6). (YYMMDD)
11> MONTH-OF-PUBLICATION	PIC XXX. (JAN/FEB/.../DEC)
12> NO-OF-COPIES	PIC 99.
13> BOOK-LANGUAGE	PIC X(10).
14> SUGGESTED-BY-MEM-NO	PIC 9(4).
15> SUGGESTED-BY-MEM-NAME	PIC X(30).

The purchase orders will be sent using the word processor or the package can also include it.

4.14 Transactions file

The structure of this file will depend upon the type of the transaction we will be having viz.

01> IB - Issue of a book/non-book,
02> RB - Receiving a book/non-book,
03> AB - Addition of a book/non-book,
04> DB - Deletion of a book/non-book(STATUS-OF-THE-BOOK = 2(LOST)),
05> UB - Updation of a book/non-book,
06> AM - Addition of a member,
07> DM - Deletion of a member,
08> UM - Updation of a member,
09> RT - Reservation of a book/non-book,
10> SP - Suggestion to purchase a book/non-book,
11> FP - Fine payment,
12> TD - Transaction deletion,
13> PD - Physical deletion of a book(after physical inventory).

This file will be of a sequential organization. The record length and fields for each transaction will be as follows...

IB -- 1>Transaction date.
    2>Transaction type(IB).
    3>Accession number.
    4>Member number.
RB -- 1>Transaction date.
    2>Transaction type(RB).
    3>Accession number.
    4>Member number.

AB -- 1>Transaction date.
2>Transaction type(AB).
3>Data entry of the first 16 fields of the Books file and initialization for the rest of the fields or a similar corresponding data entry of non-book addition. In the Books file 29 fields will be written, whereas in the Transactions file only the first 16 fields will be written.

DB -- 1>Transaction date.
2>Transaction type(DB).
3>Accession number.
4>Member number.

UB -- 1>Transaction date.
2>Transaction type(UB).
3>First 16 fields of book/corresponding fields of non-book.

AM -- 1>Transaction date.
2>Transaction type(AM).
3>Data entry of the first 5 fields for S/T/H/N/M/P with initialization for the rest of the fields of the Members file for S/T/H/N/M/P. In the Members file all the fields will be written, whereas in the Transactions file only the first 5 fields will be written.

DM -- 1>Transaction date.
2>Transaction type(DM).
3>Member number.

UM -- 1>Transaction date.
2>Transaction type(UM).
3>Fields 1 to 5 of the Members file.

RT -- 1>Transaction date.
2>Transaction type(RT).
3>Accession number.
4>Member number.

SP -- 1>Transaction date.
2>Transaction type(SP).
3>Data entry of all the 15 fields of the Purchase file.

FP -- 1>Transaction date.
2>Transaction type(FP).
3>Member number.
4>Accession number.
5>Amount of fine.
6>Date-of-payment.

TD -- 1>Transaction date.
2>Transaction type(TD).
3>Accession number.
4>Member number.
5>Old transaction type.

PD -- 1>Transaction date.
2>Transaction type(PD).
3>Accession number.

4.15 Trans-started file

In our main scheme of programs, we will be writing the following 2 main types of programs viz.

1> **Data Entry Programs** : These programs will be required because the initial data will be huge and validation / updation of initial data will be very much required, and

2> **Processing Programs** : Using these programs we will be able to process the transactions as well as get the various reports and queries.

But once the processing starts we should not be allowed to use the Data Entry programs. To achieve this we create externally one file having just one field(through EDLIN / vi/ WS) showing "Y" indicating that the processing has started. This has to be done by the systems analyst/programmer and this should not be made known to the operators. The 01 level of this file will contain only one field as follows...

```
   01 TRANS-STARTED  PIC X.
```

The organization of this file will be line sequential.

4.16 Last-day file

In this scheme, once we begin a day say for 01/07/93, we can post transactions till this date is open e.g. we can even post the transactions on 03/07/93 if 02/07/93 was a Sunday and there was some problem in the machine and we could not post the transactions on 01/07/93, but we accumulated the transactions. Also there may be the case that library starts at 10:15 a.m., they do the begin day, post some transactions, library may have lunch break, for this they stop the machine, come after lunch, restart the machine and want to post transactions. This can very well be taken care of because the Transaction Entry program will check whether the begin day for that date was done, and the close day is not done.

But once a person wants to begin a day without closing the previous day to play some tricks, we must be able to prevent this i.e. we must not allow a person to begin a new day unless the previous day is closed. For this we require one file which will contain only one field - date, which will be updated once a Close Day program for a particular date is run. The record in this file will be

```
   01   PREVIOUS-DATE     PIC X(6). (YYMMDD).
```

The organization of this file will be sequential.

4.17 Control file

Our scheme of programs will be...

1> Begin day for a particular date,
2> Post all the transactions,
3> Enquire/print the status reports as and when required, and
4> At the end of the day we must perform the following five operations...
    a> Day Book printing : This will list all the transactions of the date.
    b> Overdues printing : This will find all the overdue books,write letters and execute the Fine Calculations program.
    c> Control report printing : This will print counters for all the types of transactions. The reports in a> and c> have to be manually checked by the operators.
    d> Take the back-up of all the seven files.
    e> Close the day.

If the functions a>, b>, c>, d>, and e> are not performed then the system will not allow to start a new day.

To achieve this we need a Control file which should be an indexed file with date as the primary key. The fields of this file will be...

```
01> DATE                       PIC X(6). (YYMMDD)(Begin day date)
02> FINE-CALCULATIONS-DONE     PIC X.
03> DAY-BOOK-PRINTED           PIC X.
04> OVERDUES-REPORT-PRINTED    PIC X.
05> END-OF-DAY-REPORT-PRINTED  PIC X.
06> BACK-UP-TAKEN              PIC X.
07> CLOSE-DAY-DONE             PIC X.
08> 13 COUNTERS EACH OF PIC 99 FOR ALL THE DIFFERENT TYPES OF TRANSACTIONS.
```

4.2 Program Specifications

4.21 Program 1 : Password program

To prevent the unauthorized use of the system we introduce a password system. A person knowing the password will only be able to use the system. Instead of giving a fixed password, I suggest a password which will be changing depending upon the logic.

Logic : Get the system date and time from the terminal, move them to some fields in the working-storage section, display the fields from the working-storage section on the screen. Suppose the date is ddmmyy and the time is hhmmss, then the system(our computer program) will calculate the password as mm(of date) + mm(of time) + 19. A person wanting to access the system should also calculate the password as the above scheme and enter the password. In case both the passwords match then the system will chain to the First Menu program, else it will give three chances to enter the correct password. If the user could not give the correct password in the three chances, then the system will give the message "You are not a authorized user...stop the machine" and the system will halt.

The advantage of this scheme is that the password will be changing at every minute and a person who knows the logic can only access the machine. If the logic becomes known to some unauthorized person, then the program should be modified and another logic should be implemented.

Program 2 : First Menu

Once the valid password is given the system will display the first menu containing the five options viz.

1. Data Entry
2. Transaction Processing
3. Enquiry
4. Report Printing
5. Exit

4.22 Data Entry Module : 2 sub Modules

Program 1 : Members' Data Entry

Write a program for data entry (AA - Add After, IN - Enquire, UP - Update, DL - Delete, LI - List on printer, EX - Exit) for the Members file. Fields up to field no. 5 should be used for data entry(AA and UP) rest of the fields should be present in the file with initial values which will get updated in the transaction processing. Special care should be taken in "AA" to check the duplication of a member no. and the primary key should not be allowed to be updated.

Options for "IN" and "LI"

a> On a member number,
b> On a member name(partial name should be allowed),
c> Department wise and for a particular department,
c> Members' type wise and for a particular member type,
d> From the starting member no. to the ending member number.

For "IN" and "LI" all the fields should be used.

Program 2 : Book/Non-book Data Entry

Write a program for data entry (AA - Add After, IN - Enquire, UP - Update, DL - Delete, LI - List on printer, EX - Exit) for Books file. Fields up to field no. 16 should be used for data entry(AA and UP). All the other fields should be present in the file with the initial values which will get updated in the transaction processing. Care should be taken in "AA" to check the duplication of a member no. and the primary key should not be allowed to be updated.

Options for "IN" and "LI"

a> On an accession number,
b> On a title(partial title should be allowed),
c> Classification code wise,
d> Department wise and for a particular department,
e> Author wise and for a particular author(partial name should be allowed),
f> Publisher wise and for a particular publisher,
g> From a starting accession number to an ending accession number,
h> Title/author starting with a given alphabet or alphabetically.

For "IN" and "LI" all the fields should be used.

The "EX" option of both the above programs should chain to the First Menu program. Once the data entry is over, the entire list of books and members should be taken, checked for it's correctness, corrected if necessary, and the entire list of the corrected books and members should be taken. These two lists should be preserved as this will be the starting position of the computerized system. If necessary, the physical inventory of the books should be done, and should be checked with the corrected list of the books. Once the processing starts the computer will take care of the logical inventory.

4.23 Transaction Processing Module : 11 sub Modules

The following is the list of the modules...

01> Begin Day,
02> Transactions Entry,
03> Printing Suggestions,
04> Day Book Printing,
05> Overdues Printing,
06> Fine Calculations,
07> End-of-day Report,
08> Close Day,
09> Enquire a Transaction,
10> Back-up program,
11> Reserved Books Received.

Program 1 : Begin Day

This program will ask the member to enter a date for which he wishes to post transactions. The program will search this date in the Control file. In case an entry is found, it will give an error message "Begin day is already done",and the program will exit. In case an entry is not found in the Control file, the program will compare this date with the date in the Last-day file. In case the date in the Last-day file is smaller, it will proceed further, otherwise it will give the error message... "The begin day date has to be greater than the previous date processed" and the program will exit. Suppose the date in the Last-day file is smaller then the program will look into the Control file for this last date and check whether the close day function is done for this date or not. If not it will give the error message "Close day for previous date is not done...do it first" and the program will not allow to do a begin day. In case the close day is done for the previous date, the "Begin day" will be successful and an entry in the Control file will be made for this date.

Program 2 : Transactions Entry

This program will ask the member to enter a date of entering transactions. This date will be searched in the Control file. In case an entry is not found, an error message "Begin day is not done...do it first or enter another date" will be displayed. In case an entry is found, the program will allow us to have the following 14 options. Before we start processing any transaction the Control file record for the date will be read.

IB : will check whether the book is available in the Books file and can be issued, check whether the member has got library tickets available i.e. NO-OF-BOOKS-GOT-ISSUED <2 for S, 8 for M, 10 for T, 12 for H, 15 for P, and 5 for N. In case both the conditions are satisfied update the status of the book to 1 and corresponding fields of the Books file. Also add 1 to NO-OF-BOOKS-GOT-ISSUED and update the corresponding fields of the Members file. Add 1 to the IB counter in the Control file. Also write the transaction in the Transactions file in the "IB" format. The reservation made, if any, will be canceled. If more than one reservations were there then the latest reservation will be retained, and the other will be deleted.

RB : will check whether the book is getting received in time or not by checking the field FINE-AMOUNT in the Books file. In case the member has to pay the fine, the operator will ask him to do so and tell the fine amount and exit from this program. In case the fine is not to be paid or the member has paid the fine or the member has got a special permission, subtract 1 from NO-OF-BOOKS-GOT-ISSUED and update the relevant fields of the Members file. Change the status of the book to 0 and update the relevant fields of the Books file. Write the transaction in the Transactions file in the "RB" format. Add 1 to the RB counter in the Control file.

AB : will get the details of the book from the terminal for the fields 1 to 16 and write them in the Books file with the initialization of the rest of the fields of the Books file, so also write the transaction in the Transactions file in the "AB" format. Add 1 to the AB counter in the Control file.

DB : will check the Books file, whether for this book the FINE-AMOUNT is zero or not. In case the FINE-AMOUNT is not zero, the operator will ask the concerned member to pay the fine. In case the FINE-AMOUNT = 0, delete the book from the Books file (change the STATUS-OF-THE-BOOK to 2 and initialize the relevant fields of the Books file). Subtract 1 from NO-OF-BOOKS-GOT-ISSUED and update the relevant fields of the Members file, add the transaction in the Transactions file in the "DB" format. Add 1 to the DB counter in the Control file.

UB : will call the accession no. and write the existing fields from field no. 1 to 16 in the Transactions file in the "UB" format. Get the new information for the field nos. 2 to 16, and update the Books file. Add the transaction(updated fields 1 to 16) in the Transactions file in the "UB" format. Add 1 to the UB counter in the Control file.

AM : will get the details of the member from the terminal for the fields 1 to 5 of the Members file and write them in the Members file with the initialization for the rest of the fields of the Members file, so also write the transaction in the Transactions file in the "AM" format. Add 1 to the AM counter in the Control file.

DM : will check the Members file for NO-OF-BOOKS-GOT-ISSUED = 0. If this is not = 0 then the member will be asked to return the book/books, pay the fine amount if required, otherwise the member will be deleted from the Members file. Add the transaction in the Transactions file in the "DM" format. Add 1 to the DM counter in the Control file.

UM : will call the member number and write the existing fields from the field no. 1 to the field no. 5 in the Transaction file in the "UM" format. Get the new information for the fields 2 to 5, update the Members file, add transaction in the Transactions file(updated fields 1 to 5) in the UM format. Add 1 to the "UM" counter in the Control file.

RT : will get the accession no.and the member number.In case the STATUS-OF-THE-BOOK = 1, put member no. in the reservation field, and update the Books file. Write the transaction in the Transactions file in "RT" format. There is a provision of reservation, for only 2 members. Add 1 to the RT counter in the Control file.
FP : will call the member no., check whether the amount of fine in the account's slip = FINE-TO-BE-PAID, if yes add FINE-TO-BE-PAID to FINE-ALREADY-PAID, initialize FINE-TO-BE-PAID to zero. Before this store the transaction in the Transactions file in the "FP" format, update the related fields of the accession no.'s record for which the fine was paid(set FINE-PAID = "Y") . Add 1 to the FP counter in the Control file.
SP : will do the data entry for all the 15 fields of the Purchase file, store it in the Purchase file. Add the transaction in the Transactions file. Add 1 to the SP counter in the Control file.
TD : will get the transaction date, accession no., old transaction type. Read the Transactions file for the above information. Write the entire transaction in the Transactions file in the "TD" form. Undo the effect of the old transaction and delete the transaction from the Transactions file. Add 1 to the TD counter in the Control file.
PD : will get the accession no. to be physically deleted. Check the STATUS-OF-BOOK = 2 and FINE-PAID = "Y", if yes delete the book from the Books file. Write the transaction in the Transactions file in the "PD" form. Add 1 to the PD counter in the Control file.
EX : will exit from this program and will chain to the First Menu program.

Program 3 : Printing Suggestions

This program will be used once in 15 days, in which all the fields of the Purchase file will be printed and the decision on this will be taken manually. The file will be purged after the printing or will be deleted.

Program 4 : Day Book Printing

This program will print all the transactions of a particular date on the printer in a suitable format and Control file flag will be set. All the transactions of the date should be checked manually with this report.

Program 5 : Overdues Printing

Here the Books file will be searched and all the student members for whom books are overdue, a reminder will be sent and a flag OVERDUE-LETTER-SENT will be set and Control file flag will be set. The letter will be sent only once.

Program 6 : Fine Calculations

This program will calculate fines for all overdue books and update Members file and the Books file appropriately and will set a Control file flag. All the "book bank" books will be left from calculations.

Program 7 : End-of-day Report

This program will print all the counters from the Control file for a particular date. This report has to be manually verified with all the transactions of the date. A Control file flag will be set.

Program 8 : Close Day

This program will ask for a date. Search the Last-day file for this date. In case the date is already closed, it will give an error message. If not it will check for 1> whether the Fine Calculations is done, 2> Day Book is printed, 3>Overdues report is printed, 4> End-of-day report is printed, 5> Backup is taken.

In case all the above are true it will allow us to close the file by writing this date into the Last-day file and setting the CLOSE-DAY-DONE flag in the Control file.

Program 9 : Enquire a Transaction

This program will ask the transaction date, transaction type, accession no., member no. and display the transaction on the terminal.

Program 10 : Back-up program

Here all the seven files will be copied on tapes/floppies. A son, father, grand father concept will be followed and Control file flag will be set. The back-up has to be taken daily.

Program 11 : Reserved Books Received
This program will be executed once in three days. It will read the Books file, check whether there is a reservation of the book and the book is available, if yes print all such books in a suitable format. This report will be put on the notice board for the convenience of the members. The report is required because once a member takes a certain book, he can keep it for 15 days. But he can return it on any day within 15 days and if there is a reservation for the book, the concerned member should be intimated by means of a notice.

All the programs 1 to 11 will chain to the First Menu program.

4.24 Enquiry module : Enquiry program

The options available will be as that of "IN" and "LI" routines of the Data Entry programs for Books and Members files. The only difference is that in those Data Entry programs only restricted fields were used. Here all the fields will be used. This module will chain to the First Menu program.

4.25 Report Printing Module : 4 sub Modules

Program 1 : Statement Printing program
Here we will have three choices viz. for a single member or for all the members or for a particular member type. In all the cases the program will read the Members file and print on the terminal or the printer(as desired) the entire contents of the record/records in a suitable format.

Program 2 : Generation of Usage Statistics of Books
This program will search all the issues made till date, sort them title wise and give a report of which book has got issued how many times.

Program 3 : Printing Catalogs
This program will sort the Books file department wise, followed by title wise and print department wise catalogs with proper headings.

Program 4 : Computer Inventory of Books
This program will be run once, at the end of the year, to get the list of the books in the computer to be checked by physical inventory.

All the above four programs will chain to the First Menu program.

5.0 References

1> Fundamentals of Systems Analysis
J. F. Gerald and A. F. Gerald.
2> Information Systems for modern management
R. G. Murdick and J. E. Ross
3> Computer Education for you
December 1984 and July 1985 issues
4> WIPRO's Banking accounting system
- a brochure
5> WIPRO's Library Information System
- a brochure

6.0 Acknowledgements

The author is highly thankful to Prof. B. M. Thakare, Principal, B. N. College of Engineering, Pusad and Prof. D. J. Das, H.O.D., Computer Science and Engineering, B. N. College of Engineering, Pusad for making the facilities available for this work.

The author also expresses his heart felt thanks to Mr. S. J. Doiphode and Mr. R. B. Patil of B. N. College of Engineering, Pusad for their checking of the manuscript.

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems',
Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 253-264

A Review of Japanese Approaches to Software Production

M. Bush & D. Preston

School of Computing, Information Systems & Mathematics,
South Bank University, Borough Road, London SE1 0AA, UK

ABSTRACT

This paper provides an overview of the Japanese software industry, outlining typical approaches to software production in the leading-edge companies. It explains the layered nature of the Japanese software industry, in which software production techniques vary considerably from 'top' to 'bottom'.

Most analysts are agreed that differences in culture are the main reason for differences in industrial practices between Japan and 'the West'; in this paper we review the key differences and highlight their consequences in terms of software production.

THE SOFTWARE INDUSTRY IN JAPAN

Although the software industry in Japan has not yet reached the high levels of success attained by other industries, there are signs are that it may do so in the near future. As in other areas of Japanese industry, it is the handful of major producers that are very much leading the way, with smaller companies lagging somewhat behind.

Industry Structure

The software industry in Japan has a layered structure, as shown in figure 1. At the 'top' are the very large companies such as Japan Air Lines, as well as the major computer manufacturers (Hitachi, Fujitsu, Mitsubishi, NEC and Toshiba), who produce much of their own software. Then there are about 600 companies which are affiliated to one or other parent company that acts as the sole or principal customer. Many of these are 'spin-offs' of the parent company, set up specifically to undertake software development and maintenance. In addition there are about 50 truly independent software houses that do in-house development for a range of clients.

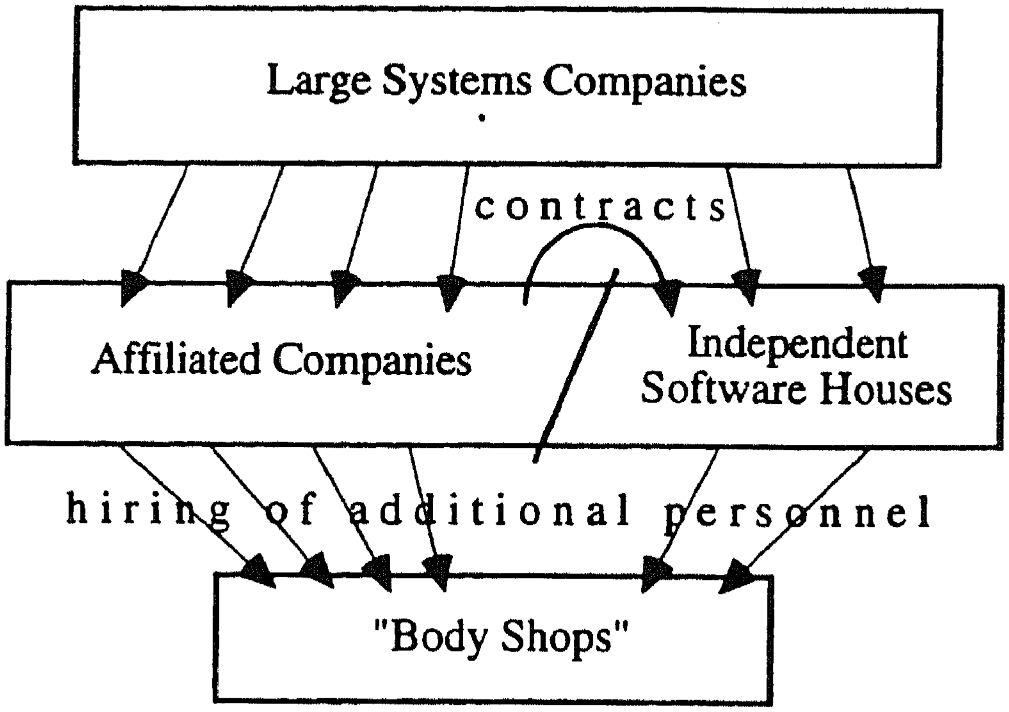


Figure 1. Layered Structure of the Japanese Software Industry

At the 'bottom', there are 5,000 or more small independent companies that act as body shops, hiring out engineers to work on projects based at customer sites. Contracts originating from the major customers (such as computer manufacturers, banks and transport operators) often filter down through their own affiliated companies to independent software houses in the form of sub-contracts. These may in turn hire additional body shop personnel to help them out on a short term basis.

Software development practices vary a great deal throughout the industry. The most sophisticated software producers are the large computer manufacturers and their affiliated companies. In addition, NTT, the recently privatised national telecommunications company (and now the richest company in the world by share value), is amongst the leaders. These are the companies which have pioneered the development of the so-called "software factories". These high-profile companies traditionally have first pick of new recruits, and because very few of their employees ever leave to join other companies there is a considerable gap in ability and effectiveness between these leading software producers and the rest. An article by Matsubara in IEEE Software [Matsubara93] argues that "the technological gap between the sunny and dark sides is widening rapidly".

This situation is exacerbated by the fact that individual esteem within society is to a large extent governed by the prestige of the individual's employer; a software engineer who is employed by a third rate company is considered (almost by definition) to be a third rate software engineer.

JAPANESE HISTORY AND CULTURE; KEY POINTS

Numerous studies in recent years have attempted to analyse the way Japanese industry operates and to draw lessons from its success. While the underlying reasons for it are many and varied, the conclusions of such studies have invariably highlighted cultural differences between Japan and 'the West' as being of vital importance. ('The West' here means western Europe and north America, and other nations such as Canada and Australia which have a similar cultural heritage.)

In order to understand the Japanese way of working it is essential to appreciate the basis of their culture, which gives rise, generally speaking, to rather different attitudes towards work than are usual in the West. Whilst it is always important to regard generalisations with a degree of scepticism, they are reasonably dependable when applied to Japan because the 120 million or so Japanese are a fairly uniform people.

Historical Influences

Aspects of modern Japanese life can be traced back to the *Edo* period which began in 1637, when Japan was almost completely cut off from the rest of the world for over 200 years. The tyrannical rulers forbade anyone to enter or leave Japan, causing the country to be excluded from international debate and progress in areas such as science and technology. Society in Japan was highly organised during this period, with clearly defined classes - nobility, military (Samurai), farmers, craftsmen and merchants. Personal conduct and even dress was dictated by detailed laws and regulations, which were brutally enforced by the Samurai. The masses learned to obey orders without question, and it seems that a willingness to follow orders still remains in the national psyche.

Japan's isolationism was abruptly brought to an end in 1853 when a powerful fleet of the US navy paraded itself invincibly on Japan's shore-lines and demanded that Japan opened its borders to trade. Lack of cooperation by the Japanese led to a three-day bombardment by British, American, French and Dutch ships in 1864, which persuaded the Japanese once and for all that they needed to open up and modernise.

Although the Samurai were harsh rulers, compared to earlier times the Edo period provided an atmosphere of relative peace and stability, in which the merchants were able to flourish. Their influence within society steadily grew, and they effectively laid the foundations for Japan's future commercial success. The pace of development once Japan opened up was astonishing. Within 50 years Japan had transformed itself from the status of a feudal agricultural nation to one of the world's most powerful and dynamic democratic societies, with a modern army and navy, industry of all kinds, a network of railways and a parliament.

However, the newly awakened national spirit, and a growing frustration brought on by the unwillingness of its Asian neighbours to buy its industrial products, led Japan into a succession of conflicts and conquests of its neighbours. Japan's aggression finally led to their involvement in the second world war, in which much of Japan was destroyed and Hiroshima and Nagasaki became the first and only (to this day) cities to have ever been devastated by atomic bombs.

This was in fact the very first time in recorded history that Japan had been conquered by an invader. Although Japan suffered severe damage and economic set-back, the defeat precipitated a second industrial transformation in Japan which, like the first one 80 or so years earlier, has resulted in quite amazing industrial and economic growth. Japan has now reached the status of economic superpower. Its per capita assets are three times those of the USA and four times those of the UK. Personal savings rates are currently, on average, around 14% of income; the highest in the world.

Industry

Undoubtedly an important force in Japan has been, as in the former West Germany, a national desire to regain status after the second world war. At least this has been the outlook of both countries' leaders, if not their whole populations. Since the war both countries have focused their efforts on building up their commercial industries; military spending was suppressed by conditions imposed by the allies at the end of the war. This has given them a considerable advantage over countries such as the UK, France and the US which have maintained much higher levels of military spending. For example, Japanese companies now operate using very high technology platforms as a result of on-going investment. More industrial robots and office-automation equipment are to be found in Japan than in any other country.

Research and development in Japan tends to be company-led, rather than government-led. The large companies often work together on major R&D projects. They also cooperate on the development of standards which are subsequently adopted nationally. In the West things tend to work the other way around; standards are proposed by government bodies often working together with academic institutions, and only then are companies encouraged to adopt them.

Industrial success for the Japanese has come largely from their long-standing commitment to, and now their reputation for, high quality. This commitment can be traced back to the early years after the war when several industrial 'gurus' from the USA went to Japan to act as advisers in the rebuilding of Japanese industry. The most famous of these was W. Edwards Deming, whose management principles are still much honoured.

Many companies in Japan operate as part of 'business families'. Before the second world war these were known as *Zaibatsu*. A typical Zaibatsu consisted of a tightly coupled collection of key manufacturing and service organisations under the influence of one major share-holding company. The chairmen of the largest Zaibatsu were very influential members of society. However, all this changed after the war. The Zaibatsu were dissolved and new laws were put into place effectively preventing individual members of society acquiring too much personal wealth and influence.

Instead, some of the groups of companies which had originally formed Zaibatsu got together to form business families of a rather different kind. The idea was still that these companies would support each other as much as possible through trade and servicing agreements, but now the companies were free to operate in their own best interests. Bonds between the companies were usually reinforced by a small amount of mutual share-holding.

Many of these business families are still in operation, and indeed some are very successful. Often such a family will include a bank, a computer/electronics company, a pharmaceutical company and other key manufacturing and service organisations. Well known examples are the Mitsubishi, Sumitomo, Furukawa and Mitsui groups. Decision making within these families is distributed amongst the member companies. By contrast, in the Zaibatsu the decision making power was centralised.

In other cases smaller companies cluster around one giant; this is the case with Toyota, Honda and Hitachi. Such clusters are known as *Keiretsu*. Although similar in structure to Zaibatsu, important differences are the fact that the central giant is always a major producer of goods or services, and itself must be under the directorship of a board of directors. Business empires of the kind controlled by Robert Maxwell or Richard Branson are not to be found in post-war Japan.

In general, Japanese companies have long-term rather than short-term ambitions, and are therefore inclined towards increasing their market share rather than increasing short-term profits. It is this attitude which often leads to Japanese products being exported at very low profit margins, much to the distaste of foreign governments who see these cheap imports as a threat to their national manufacturing industries. In many cases Japanese products are actually more expensive to buy in Japan than they are in some other countries. Virtually all the profits which are made are invested in areas such as research and development, and in improving the manufacturing process. Most Japanese companies seem to be continually and self-consciously striving for improvement.

Despite several recent financial scandals and allegations of corruption, most business in Japan is sincere, open and ethical. There is usually a good deal of trust involved in business deals, with few litigation proceedings. Consequently the Japanese have a rather relaxed attitude towards written contracts. This has particular implications for the software industry, as we shall see.

Employer/Employee Relationships & Attitudes

Although Japan is now in many senses a 'Westernised' society, with countless discos, fast food restaurants and amusement arcades, their way of life, and to a certain extent their outlook on life, is still distinctly different. Whereas the principles and philosophies of the West are derived from such influences as Judeo-Christianity, Greek rational thought and Roman administrative systems, those of Japan stem from Buddhism, Shintoism and Confucianism. One of the main differences which arises is that whereas, deep down, most Westerners regard themselves as true individuals who 'only live once', most Japanese tend to regard themselves fundamentally as part of the society in which they live, which, as in Confucianism, is thought to be the greatest good.

Generally speaking, the Japanese can be said to have a tendency towards:

- respect for elders, and for seniority within society
- respect for ethics and morals
- willingness to conform
- perfectionism
- diligence
- team working
- competitiveness
- long-term thinking.

The picture of the stereo-typical Japanese male industrial worker is well known. He is devoted to one company, never having worked for any other, and lives with his family in the same neighbourhood as many of his colleagues, perhaps in a company-owned apartment. The school his children attend is probably also supported by the company. His family often joins those of his colleagues on company holiday excursions. In these respects Japanese society is almost communist in nature.

The situation is changing, though. Whereas before the second world war many Japanese may genuinely have conducted their lives for the good of society as they saw it, now they are no doubt mostly motivated simply by wanting to do well for themselves and their immediate families. To that extent they have become more Westernised; the ideals of Buddhism and Confucianism have definitely begun to wane. A book written recently by leading politician Ichiri Ozawa ("Blueprints for Building a New Japan") identifies excessive centralisation and regulation, the dilemma of ageing, sexism, servility to employers and overwork as Japan's severest current social and economic problems [Time93].

Although personal income is roughly the same as it is in the USA or the former West Germany, Japanese employees have to work an average of 2,160 hours a year to earn it, compared with 1,980 in the USA and only 1,640 in the former West Germany. Indeed, the Japanese National Defence Council for Victims of *Karoshi* (death from overwork) estimates that as many as 10,000 individuals actually die each year as a direct result of overwork. Compensation payouts to families of victims are rare, although claims are sometimes made. In 1992 about 600 cases are thought to have been filed. A settlement announced in September 1993 by Japan Tabacco Inc. - of $370,000 - to the family of a worker who died in 1984 after having been forced to work more than 400 hours was the biggest single payment agreed to date [Herald93].

Japan has a long tradition of excellence in craftsmanship, and most Japanese workers do seem to take considerable pride in their work. As customers they also expect the best; the discerning nature of the average customer in Japan was arguably the original motivating force which caused Japanese companies to focus on producing products of high quality.

Consequences of Life-Time Employment

The major employers in Japan operate on the principle of life-time employment, and this has a number of distinct advantages with respect to the relatively high staff turnover rates common in the West. For example:

- workers are likely to take more care over their work, perhaps because they feel a stronger sense of loyalty to the company, and because they are more likely to still be around in the future to face the consequences of any shoddy work
- communication and cooperation between co-workers is likely to be more effective since they will have worked together for many more years, on average, and are quite likely to be social acquaintances
- administrative and other systems work more smoothly on the whole because everyone is more familiar with the procedures, through having lived with them for longer.

The most successful Japanese companies appear to operate at a level of harmony which few Western companies seem to be able to match. No doubt the factors just mentioned have a lot to do with this. By contrast, there is often tension within many companies in the West caused by conflict between different departments, and by repeated changes in company structure brought about by changes in upper management. Newly appointed managers in the West too often want to 'make their mark' by reorganising existing structures or procedures. This usually creates a considerable amount of upheaval, and although probably well intentioned, may well cause matters to get worse rather than better. Their replacements in turn create more upheaval, and those lower down the company hierarchy who may have lived through several such upheavals can easily become disillusioned.

In return for a high level of commitment from its employees, most Japanese companies offer extremely good employee training and care programmes. In times of difficulty Japanese companies make every effort to re-train and re-deploy staff if at all possible, rather than making staff redundant. A good example of this is Nippon Steel Corporation, which in recent years has re-trained around 800 steel workers, with an average age of 43, to become software engineers. The motivation was simple; on one hand the steel industry was in a slump and on the other hand they were in need of software engineers. In similar circumstances, Western companies would be much more likely to make 800 workers redundant and independently recruit 800 qualified software engineers.

Volunteer steel workers at Nippon Steel Corporation underwent one year of full time training, after which they became junior programmers. The drop out rate was only about 10%, and apparently most of them have become compètent software engineers. It is claimed that part of the reason for their success is that they are well used to working closely together in tight knit teams, and that this is an important quality when it comes to software engineering.

This kind of training is not uncommon. It is quite normal for companies to provide new recruits with several months full time training before they are expected to produce any real work. No European or American company can afford this kind of training because staff turnover is so high by comparison; in fact turnover is particularly high for employees in their early working lives. It is also normal in Japan for newly appointed managers to receive two or more weeks of full time intensive training in management techniques, with additional intensive training each time they move a step up the management ladder.

Figures show that the average Japanese software engineer actually attends 13 times as many courses as the average American software engineer, although most of these courses are in-house and by all accounts they can be somewhat ritualised.

TYPICAL SOFTWARE PRODUCTION TECHNIQUES

On the whole, there does not seem to be a major difference between software development practices in Japan, Europe and the USA, although minor differences certainly exist. For one thing, the low turnover rate in Japan has allowed them to rely less on documentation, since the original designers are usually still around when it comes to maintenance. Another factor here is that, due to the complexities of the written language, Japanese word processors have only been developed relatively recently.

As in other areas of business, there is also less emphasis on written contracts in Japan; Japanese companies are used to working on the basis of trust. Litigation proceedings are seen as bringing shame upon both supplier and customer. Consequently there is nothing like the interest in third party testing or certification that there is in the West where, particularly with the opening up of free trade between European countries, certification is seen as providing some kind of guarantee of product quality. The principal motivation for Japanese companies to adopt standards such as ISO9000 comes from their desire to compete in world markets where such standards are generally seen as playing an increasingly important role.

Although the Japanese have considerable experience at producing very complex systems such as large-scale telecommunication systems and sophisticated computer operating systems, they have rarely pushed forward the state-of-the-art of software technology. Also, rather than develop new production techniques, the Japanese have very often chosen to copy and refine the most promising approaches used in the West. According to Cusumano, the notion of a software factory originated not in Japan but in the USA in the 1960's [Cusumano91]. IBM, SDC and TRW in the USA all tried out software factory concepts, although SDC were the only ones to use the 'factory' label. However, while SDC allowed its factory experiment to come to an end in 1979 after teething problems, the Japanese factories, through greater perseverance, appear to be going from strength to strength.

One reason why the Japanese were so keen to copy was that IBM had over 50% of the world market share of computers in the 1970's. Fujitsu and Hitachi in particular decided that their best strategy would be to produce machines with IBM-compatible operating systems in the hope that users would see fit to migrate their existing software. This led to drawn out legal disputes with IBM. A settlement was finally reached between Fujitsu and IBM in 1988, lasting ten years, which essentially gave Fujitsu the right to continue copying IBM machines after paying a lump sum of several hundred million dollars, plus an annual licensing fee.

The idea that Japanese companies have become successful simply by copying American products and concepts is, however, far from the whole truth. Companies such as NEC and Toshiba chose to take a different path and have developed entirely new operating systems. In fact there have been numerous American visitors to the major Japanese software producers since the 1970's who have tried to emulate at home aspects of what they saw there. Of course, a major reason why it is the Japanese who have been able to do most of the copying is that they are so familiar with English. If only Westerners could read Japanese books, reports and research papers with the fluency that many Japanese can read English ones, the situation would surely be quite different.

In 1986 the total revenue for software in Japan was $5 billion, compared with $19 billion in the USA. However, as much as 96% of the Japanese software was designed for integrated software and hardware systems, in comparison with nearly 60% of the US software which consisted of mass market packages. This may go some way to account for one of the major apparent differences between Japan and the West, which is that the large Japanese firms on the whole are more concerned with process improvement while Western firms are often more concerned with product innovation. Japanese software factories try wherever possible to tackle projects which are similar to ones they have tackled before. Novel projects are usually tackled outside the factory environment. This may also account for the apparently higher figures for reused software in Japanese products.

Japanese software products on the whole seem to be more reliable than those produced in the West - figures vary, but typically indicate that Japanese software products have something like 50% fewer faults. One consequence of this is that maintenance costs are greatly reduced. Demand for software within the Japanese domestic market continues to exceed the supply capabilities of the software producers, and this has been one of the main reasons why the Japanese have not exported much software or bid frequently for overseas contracts. It is dangerous to assume that, because they do not currently export much software, their software industry is immature. On the other hand, producing software products for foreign markets is considerably harder than producing the same product for the home market because of the language barrier.

The Software Factory Approach

The term "software factory" became popular in Japan in the 1970's, and was adopted by several Japanese companies. Hitachi were the first to use the term, in 1969. Two factors which prompted Hitachi to set up their software factory were a severe shortage of skilled software engineers and a poor past reputation for quality. Hitachi set two goals for their newly created software factory:

- to improve productivity and reliability through process standardisation and control
- to transform software production from an unstructured service to one which could deliver products with a guaranteed level of quality.

These goals are just as appropriate today as they were then. In order to achieve them, a software factory must have a standardised life-cycle model and project management procedures, and a uniform development environment utilising CASE (Computer Aided Software Engineering) tools as far as possible. It must also have a well developed data collection and analysis programme so that quality and productivity trends can be tracked and new working methods properly assessed. The notion of a software factory also places an emphasis on the use of reusable code modules.

Although strictly speaking the term 'software factory' refers to a concept rather than a physical building, often it does refer to a particular location where lots of people are engaged in software production. Typically, a software factory consists of large open plan offices with 100 or more software developers in each. Workspace is often very limited, particularly in urban areas where land prices are high. Each engineer normally has a small area of desk space, a computer terminal, and shelf and some filing space. Desks are often set out in rows, with group leaders in a position to see all the members of their group. It is also common for the department manager and his or her deputy to sit in partitioned rooms in a corner of the office.

Of course, simply adopting the factory title does not necessarily mean a different way of working, and, in general, the tools and techniques used in Japan (whether in software factories or not) are not very different to those used in the USA and Europe. The review or walkthrough, for example, is still the principal technique used for improving the quality of software. However, most Japanese software factories have launched genuine long-term efforts to centralise and systematise their software production.

Hitachi tried to run its software factory along the same lines as its hardware factories. As in other companies, the drive came from upper management rather than the software engineers themselves, who tended to dislike the idea at first, feeling that they were white collar workers being asked to operate along the same lines as blue collar workers. During the mid 1970's there was heavy investment in the development of automated tools for project management, design support, testing, program generation and reuse support. In 1985 Hitachi split its software operations into 'basic' and 'applied', since it realised that different tools, techniques and methods were needed for each.

In 1976 NEC established no less than 5 software factories, with between 1,250 and 2,500 software engineers in each one. Since then Toshiba, Fujitsu, NTT and Mitsubishi have all followed suit. (Interestingly though, Hitachi have recently changed the name of their software factory to 'Software Development Centre' because they feel the factory image is actually off-putting to prospective new recruits!)

The way quality assurance is normally tackled within the Japanese software industry is rather different from the way it is normally tackled in the West. In Japan every member of staff is encouraged to be involved with quality issues, whereas in the West there is often a detached quality team. Also in Japan quality assurance programmes are often applied company-wide rather than just project-wide, as is more usual in the West. Open relationships between suppliers and customers in Japan mean that in many cases customers are given a high degree of visibility into how projects are progressing. For example, customers may be shown graphs showing the trends in defect detection rates during various phases of the life-cycle; something which few Western suppliers would contemplate.

REFERENCES

[Cusumano91]	Cusumano, M., *Japan's Software Factories*, Oxford University Press, 1991
[Herald93]	*Worked to Death, but Paid for it*, International Herald Tribune, 1 Oct 1993
[Matsubara93]	Matsubara, T., *To Improve Averages, Focus On The Worst*, IEEE Software, Vol. 10, No. 3, p5, May 1993
[McQueen89]	McQueen, I., *Japan: A Travel Survival Kit, Third Edition*, Lonely Planet Publications, 1989
[Mikes70]	Mikes, G., *The Land of the Rising Yen*, Penguin Books, 1970
[Ouchi81]	Ouchi, W., *Theory Z*, Avon Books, 1981
[Time93]	*Shortest Distance Between Points*, Time International, 4 Oct 1993

Proceedings Intern. AMSE Confer. 'Signals, Data, Systems', Bangalore (India), Dec. 8-10, 1993, AMSE Press, Vol. 1, pp. 265-276

EFFICIENT SCHEDULING OF PRODUCTION AND SETUPS ON A FLEXIBLE MACHINE USING LAGRANGIAN RELAXATION

Sundar Ram Vedula
Department of
Mechanical Engineering
Indian Institute of Technology
Madras 600 036, INDIA

Y. Narahari
Department of
Computer Science and Automation
Indian Institute of Science
Bangalore 560 012, INDIA

Abstract

In this paper, we develop an algorithm to schedule jobs on a single machine in a multiclass production environment. We start with a schedule obtained by Lagrangian Relaxation [8], and modify it suitably by regrouping jobs belonging to similar classes, so that setup times during changeovers are considered. The algorithm is illustrated for a two-class, 89 job scheduling problem and results are presented for different setup times.

1. INTRODUCTION

Scheduling is one of the most important issues in the planning, management and operation of manufacturing systems. In this paper, we address the problem of setups, or switchover times incurred in a multiclass production environment. We propose a simple heuristic algorithm, based on Lagrangian Relaxation [1, 8, 9] to produce a sub-optimal schedule that effectively handles the need for switchovers between multiple job classes. The algorithm is tested for various setup times on an 89 job example.

1.1 Schedule Generation using Lagrangian Relaxation

The theory of sequencing and scheduling, more than any other area in Operations Research, is characterized by a virtually unlimited number of problem types [2, 6]. Recently, in the area of deterministic machine scheduling, scheduling algorithms based on Lagrangian Relaxation [1, 7] have been proposed and have proved to be highly effective in real-world factory situations. These Lagrangian Relaxation based algorithms have been reviewed in [5]. Luh et al. [8] presented the first of these algorithms for the case of scheduling single operation jobs on parallel identical machines. Later, the approach was extended to the case of multiple-operation jobs with precedence constraints [4]. The algorithms use a decomposable objective function such as the total weighted tardiness and obtain the solution by solving independent, job-level subproblems. All these algorithms have proved to be efficient, near-optimal procedures, with the additional feature of yielding a lower bound on the cost, which can be used as a measure of suboptimality.

1.2 The Problem of setups

The Lagrangian Relaxation (LR) technique works on the problem by decomposing it into a number of job-level subproblems, which may be solved independently. However, once setups are introduced, this advantage is lost, as the cost cannot be independently evaluated for each possible beginning time of a job. Different beginning times of the job put it in different batches depending on the location of other jobs, and this causes loss of independence and decomposability of the original problem. So the LR technique cannot handle setups directly. Also, the original problem itself is strongly np-hard [7, 8, 9, 11]. Our goal is therefore not to obtain the optimal solution. Rather, we obtain an efficient solution that includes the effect of setups and improves on an initial schedule generated by the original LR technique.

1.3 Organization of the Paper

In Section 2, we review the LR technique developed by Luh et al. [8]. In Section 3, we describe in detail the algorithm we propose to generate an efficient schedule taking setups into consideration. The input data and numerical results for an 89 job scheduling problem are illustrated in Section 4.

2. A LR TECHNIQUE FOR SCHEDULING

The basic technique that we employ in this paper is the recent one proposed by Luh et al. [8]. We give a brief description of their technique here. The method is used for the nonpreemptive discrete time scheduling of independent jobs on identical, parallel machines and is based on the Lagrangian Relaxation Algorithm [1, 7]. The technique yields a near-optimal solution and has the feature that its performance can be measured with respect to the optimum. The input variables are:

N : Number of jobs
K : Time horizon for scheduling
$w_1, \ldots w_N$: Weights (relative importance) of jobs
$t_1, \ldots, t_N$: Processing times of the jobs
$D_1, \ldots, D_N$: Due dates of the jobs
$M_1, \ldots, M_k$: Numbers of machines available at discrete time instants 1, 2, .., K, respectively.

The decision variables are $B_1, \ldots, B_N$, the beginning times for the jobs. Now, define, for $i = 1, .., N$, and $k = 1, \ldots, K$,

$s_{ik} = 1$ if job i is undergoing processing at time k
$s_{ik} = 0$ if job i is not undergoing processing at time k
C_i = Completion time of job i
$T_i = \max\{0, C_i - D_i\}$

T_i represents the tardiness of job i. The objective function of interest is

$$J = \sum_{i=1}^{N} w_i T_i$$

which is the weighted tardiness criterion. This criterion takes into account the relative importance of jobs, and the importance of meeting due dates. The scheduling problem can now be stated as follows.

$$\min_{\{B_i\}} J \text{ with } J = \sum_{i=1}^{N} w_i T_i$$

subject to capacity constraints:

$$\sum_{i=1}^{N} s_{ik} \leq M_k, \; k = 1,\ldots, K$$

and processing time requirements:

$$C_i - B_i + 1 = t_i, \; t = 1,\ldots,N$$

The problem is solved by maximizing the dual function which is obtained by using a set of multipliers to relax the capacity constraints. The dual problem is

$$\max_{\pi} L \qquad \text{with}$$

$$L = \{ -\sum \pi_k M_k + \min_{\{B_i\}} \sum_{i=1}^{N} \{w_i T_i + \sum_{i=1}^{K} \pi_k s_{ik}\}\}$$

subject to $\pi \geq 0$ and

$$C_i - B_i + 1 = t_i, \; i = 1,\ldots, N$$

The dual problem can be decomposed into N subproblems, one for each job, as follows. Given π and $i = 1,\ldots,N$, the i^{th} subproblem is

$$\min_{1 \leq B_i \leq K - t_i + 1} L_i \quad, \text{with}$$

$$L_i = w_i T_i + \sum_{k=1}^{K} \pi_k s_{ik}$$

subject to $\quad C_i - B_i + 1 = t_i$

The computational complexity of each subproblem is linear in K. Using these K solutions, the dual problem is solved using a subgradient method. The stopping criterion used in the above subgradient method does not always yield a feasible schedule, in the sense that the capacity constraints might be violated for a few time slots [8]. To construct a feasible schedule, a greedy heuristic based on the list scheduling concept is used.

An interesting feature of the above technique is that the performance of the feasible schedule obtained can be measured against that of the optimal schedule [8].

3. SCHEDULING OF PRODUCTION WITH SETUPS

For a single machine, we propose the following technique to include the effect of setups on the final schedule. Initially, we employ the LR technique as in Luh et al. [8] to generate a sequence of jobs scheduled to be produced on the given machine. Assuming n jobs, Let the obtained sequence be Jx_1, $Jx_2, \ldots, Jx_n$. This sequence is nearly optimal if setup times are neglected. However, with setups, this may prove to be highly inefficient as shown in Section 4. Our idea is to now group the jobs belonging to various classes together without upsetting the basic sequence.

3.1 Scheduling with two job classes

We illustrate our approach for the two job-class case. Looking at the schedule of jobs on the machine, we see that every job belongs to either class A or class B. We pick out all the class A jobs as they appear in the schedule and designate them Ja_1, $Ja_2, \ldots$, Ja_p and similarly the class B jobs and designate them Jb_1, $Jb_2, \ldots, Jb_q$, where p = number of class A jobs, q = number of class B jobs, and $p + q = n$.

Assuming a known number of setups k (This is discussed later) we designate k variables S_1, $S_2, \ldots$, S_k as having values corresponding to the number of jobs after which a setup takes place. Since we assume that the machine is initially idle, $S_1 = 0$ always. For example, $S_2 = 10$, $S_3 = 15$ means that 10 jobs of one class are processed, followed by a setup, and then 5 additional jobs of the other class are processed followed by a setup and so on. We first look at the schedule having equal batch sizes for class A jobs and equal batch sizes for class B jobs. The batch size for A will be $q_a = 2n_a/s$ and that for B will be $q_b = 2n_b/s$. The sequence will look like one of the following:

$$Ja_1,\ldots,Ja_p \mid Jb_1 \ldots\ldots Jb_q \mid \ldots \qquad (1)$$

$$Jb_1,\ldots,Jb_q \mid Ja_1 \ldots\ldots Ja_p \mid \ldots \qquad (2)$$

Corresponding to (1), we have

$S_1 = 0$
$S_2 = q_a$
$S_3 = S_2+q_b = q_a+q_b$
$S_4 = S_3+q_a = 2q_a+q_b$
$S_5 = S_4+q_b = 2q_a+2q_b$, and so on.

Corresponding to (2), we have

$S_1 = 0$
$S_2 = q_b$
$S_3 = S_2+q_a = q_b+q_a$
$S_4 = S_3+q_b = 2q_b+q_a$
$S_5 = S_4+q_a = 2q_b+2q_a$, and so on.

We choose sequence (1), or (2), depending on which gives a lower cost. We now see that given the values of $\{S_i\}_{i=1 \text{ to } k}$, we can generate a schedule, and compute the corresponding weighted tardiness. For describing the technique, suppose we choose sequence (1). We now decrement S_2 by 1 so the size of the first batch is q_a-1. To retain the size of the second batch, we need to decrement S_3 by 1. The batch sizes are now q_a-1, q_b, q_a. We compute the cost of this schedule. We also increment S_2 and S_3 by 1, to get the batch sizes as q_a+1, q_b, q_a. Out of the 3 possible schedules that we have, we choose the one with lowest cost and set the corresponding $\{S_i\}$ variables. To change the size of the second batch (first batch of class B jobs), we increment or decrement the values of S_3 and S_4 by 1, and again choose the schedule with lowest cost. This is repeated till we increment or decrement S_{k-1}

and S_k.

At the end of the first iteration, we have a revised schedule with a cost lower than that of the schedule with equal batch sizes, due to the manner in which the permutation of jobs is done. The iterations are now repeated in an identical manner. Gradually, the values of $\{S_i\}$ are seen to stabilise, that is, it proves unworthy to change the value of S_i in a given iteration. Once we reach a stage where no further decrease in the objective function is possible, the procedure is stopped. We have now obtained a value of the cost function that is sub-optimal with respect to the variables $S_1, S_2, \ldots, S_k$. Fixing of these values of $\{S_i\}$ automatically implies determination of the final schedule.

3.2 Choice of number of setups

The methodology proposed needs an initial value of k. For this selection, the following considerations may be used:

- If the number of setups is too low, then the tardiness contribution of delayed jobs becomes too high. An extreme case of this would be when we group all As together and all Bs together. If As are scheduled first, then the weighted tardiness of high priority jobs in class B becomes very high.

- If setups are too frequent, then most time is spent in setting up the machine rather than in processing jobs. This frequent idling of the machine contributes significantly to the tardiness of jobs scheduled later on it. The upper limit on the number of setups is the number of setups required for a schedule generated directly by the LR technique.

- The actual setup time in relation to the processing times of the various jobs gives a very good idea of how often it really is worth going in for a setup of the machine.

Based on the above, an estimate of the number of setups may be made and the algorithm may be run for two or three values around this estimate. The schedule giving the lowest cost is then

Table 1

i	w_i	t_i	D_i	cl	i	w_i	t_i	D_i	cl
1	1	5	0	A	2	1	10	15	A
3	1	5	12	A	4	1	5	13	A
5	1	5	1	A	6	1	5	-3	A
7	1	10	-2	A	8	1	4	9	B
9	1	3	9	B	10	1	5	7	A
11	1	5	8	A	12	1	10	13	A
13	1	5	19	B	14	1	5	2	A
15	1	5	7	A	16	1	10	-1	A
17	1	5	1	A	18	1	3	0	B
19	1	5	5	A	20	1	10	9	B
21	1	5	1	A	22	1	10	8	A
23	1	3	0	B	24	1	7	3	B
25	1	8	60	B	26	1	9	8	B
27	1	10	-1	A	28	1	3	3	B
29	1	10	0	A	30	1	5	1	A
31	1	5	-1	A	32	1	5	7	A
33	1	11	11	B	34	1	10	15	A
35	1	16	24	B	36	1	5	2	B
37	1	3	25	B	38	1	3	24	B
39	1	3	20	B	40	1	4	18	B
41	1	5	2	A	42	1	3	2	B
43	1	5	10	B	44	1	3	7	B
45	1	5	7	A	46	1	3	2	B
47	1	10	7	A	48	1	5	7	A
49	1	5	8	A	50	1	10	7	A
51	1	10	7	A	52	1	5	10	A
53	1	5	9	A	54	1	3	2	B
55	1	4	1	B	56	1	3	9	B
57	1	3	2	B	58	1	10	-1	A
59	1	5	10	A	60	1	20	24	B
61	1	5	5	B	62	1	4	9	B
63	1	6	7	B	64	1	5	5	B
65	1	5	10	A	66	1	3	2	B
67	1	3	7	B	68	1	15	42	B
69	1	8	10	B	70	1	16	17	B
71	1	10	20	B	72	1	3	-1	B
73	1	8	11	B	74	1	4	4	B
75	1	4	15	B	76	1	3	5	B
77	1	3	2	B	78	1	3	1	B
79	1	5	-1	A	80	1	3	12	B
81	1	3	10	B	82	1	3	9	B
83	1	10	4	A	84	1	3	2	B
85	1	12	9	B	86	1	10	1	A
87	1	10	1	A	88	1	10	1	A
89	1	10	0	A					

chosen.

3.3 Extension to Multiple Machines

Though the technique described applies to scheduling on a single machine, it can be very easily extended to schedule jobs and setups on identical, parallel machines. The LR algorithm [8] is used to generate independent schedules for each of the machines. Our methodology may then be applied to each of the machines separately treating them as single machines destined to process a set of jobs allocated to them.

4. NUMERICAL RESULTS

4.1 Scheduling Data

Experiments are carried out on an 89 job example, from [8] with suitable modifications. The data used is shown in Table 1. The columns contain respectively, the job number, weight, processing time, due date, and class. From this data, a schedule is first generated using LR without considering setups, and then appropriate regrouping is done to produce an efficient schedule with setups.

4.2 Results

Without considering setups, LR gives a schedule of cost 17726. With various setup times, the improvements to the schedule are summarized in the following table, where t_s is the setup time.

t_s	1	2	5	10	20	50	100	150	200	300
Cost of LR schedule	18606	19198	20974	23934	29854	47614	77214	106814	136414	195614
Cost of our schedule	17811	18015	18568	19338	20878	25498	32392	39294	46051	59181

For setup times of 1 and 2 (an order lesser than processing times) it is seen that the schedule with 5 setups gives the least cost. The order of batches is as in BABAB. As the setup time is increased, for an initial schedule with 5 setups, two of the setups are seen to converge into one another during the iterations, thus yielding a solution with just 3 setups. This occurs for setup times 5, 10, 20, 50 and 100. Beyond 100 (an order greater than processing times), it is seen that the schedule with just two setups (BA) comes very close to the schedule with three setups (BAB). These results are summarized as follows.

t_s	100	150	200	220	240	250	265*
Cost for 2- setup schedule	32392	39294	46051	48691	51331	52651	54631
Cost for 3- setup schedule	33181	39681	46181	43781	51381	52681	54631

* For values of t_s above this, the setups for an initial 3-setup schedule converge into one another to yield just two setups.

4.3 Conclusions

The heuristic algorithm presented above effectively modifies the schedule generated by the LR technique to include the effect of setup times and generates a new schedule. At every stage, we proceed along the hyperplane of $\{S_i\}$ values only in the direction of decreasing cost, and there is a very definite stopping criterion, that is, when the cost cannot be reduced further. Given the original schedule generated by LR, the final schedule is generated very quickly. Another feature of the algorithm is that giving a large initial number of setups causes adjacent $\{S_i\}$ values to become equal to one another, implying that 2 setups are eliminated for every such case, thus reducing the number of setups automatically.

5. SUMMARY AND FUTURE WORK

This paper presents a heuristic algorithm to schedule setups alongwith jobs in multiclass production system using an initial schedule generated by the Lagrangian Relaxation technique [8]. Significant reduction in cost is achieved, especially for higher setup times. Future work aims at extending the methodology to handle a greater number of job classes and to develop more intelligent criteria for the permutation of jobs between batches. Work is also in progress to develop algorithms based on LR to handle on-line changes in the production environment.

REFERENCES

1. M.L. Fisher, Lagrangian Relaxation Method for Solving Integer Programming Problems. Management Science, Volume 27,1981, pp.1-18.

2. S. French, Sequenching and Scheduling. Wiley, New York, 1982.

3. S.B. Gershwin, Hierarchical Flow Control : A Framework for Scheduling and Planning Discrete Events in Manufacturing Systems. Proceedings of the IEEE, Volume 77, Number 1, January 1989, pp. 195-209.

4. D.J. Hoitomt, P.B. Luh, E. Max, and K.R. Pattipati, Scheduling Jobs with Simple Precedence Constraints on Parallel Machines. Chapter 20, in: Y.C. Ho (Editor), Discrete Event Systems, IEEE Press, 1991.

5. D.J. Hoitomt, P.B. Luh, and K.R. Pattipati, A Practical Approach to Job Shop Scheduling Problems. To appear in: IEEE Transactions on Robotics and Automation, 1993.

6. E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, Sequencing and Scheduling: Algorithms and Complexity. Report No. BS-R8909, Centre for Mathematics and Computer Science, Amsterdam, November 1989.

7. D.G. Luenberger, Linear and Non-linear Programming. Second Edition: Addison-Wesley, Reading, Massachusetts, 1984.

8. P. Luh, D. Holtomt, E. Max, and K. Pattipati, Schedule Generation and Reconfiguration for Parallel Machines, IEEE Transactions on Robotics and Automation, Volume 6, Number 6, December 1990, pp. 687-696.

9. Y. Narahari and Sundar Ram Vedula, Real-World Extensions to a Production Scheduling Algorithm Based on Lagrangian Relaxation. Submitted to the IEEE International Conference on Robotics and Automation, San Diego, May 1994.

10. R. Srigopal and Y. Narahari, Deterministic Scheduling of Multiclass Manufacturing systems using Simulated Annealing and Lagrangian Relaxation (In preparation).

11. N. Viswanadham and Y. Narahari, Performance Modeling of Automated Manufacturing Systems. Prentice Hall, Englewood Cliffs, New Jersey, 1992.

ASSOCIATION FOR THE ADVANCEMENT
OF MODELLING AND SIMULATION TECHNIQUE
IN ENTERPRISES

16, Avenue de Grange Blanche, F-69160 TASSIN-LA-